PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/74, 15/62, 15/46, 15/44, 15/70, 15/31, C07K 14/14, 14/11 // C12N 1/21, C12R 1:225

A1

(11) International Publication Number: **WO 96/32487**

(43) International Publication Date: 17 October 1996 (17.10.96)

(21) International Application Number: PCT/NL96/00160

(22) International Filing Date: 11 April 1996 (11.04.96)

(30) Priority Data:
PCT/NL95/00135 11 April 1995 (11.04.95) WO
(34) Countries for which the regional or international application was filed: AT et al.

(71) Applicant *(for all designated States except US)*: NEDERLANDSE ORGANISATIE VOOR TOEGEPAST-NATUURWETENSCHAPPELIJK ONDERZOEK TNO [NL/NL]; Schoenmakerstraat 97, NL-2628 VK Delft (NL).

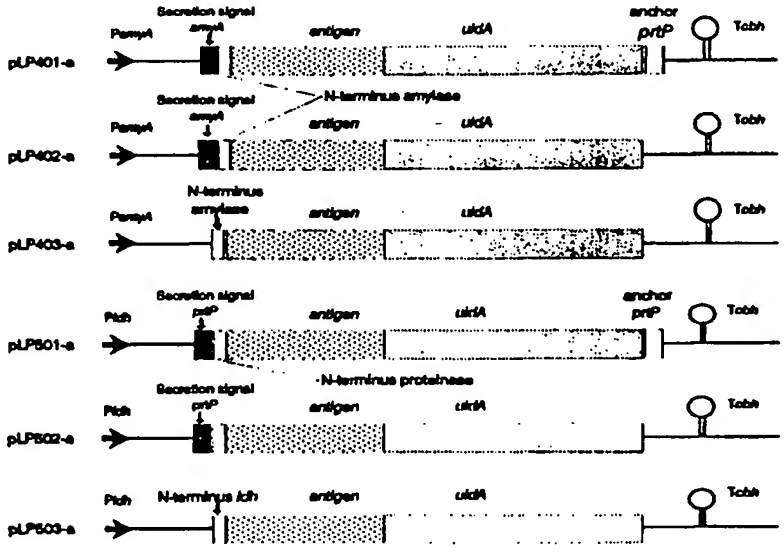
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LEER, Robert, Jan [NL/NL]; Tulpentuin 36, NL-2272 XK Voorburg (NL). POUWELS, Pieter, Hendrik [NL/NL]; Delftweg 14, NL-2289 AJ Rijswijk (NL). BOERSMA, Wilhelmus, Johannes [NL/NL]; Burgemeester Merkusstraat 7, NL-2645 NJ Delfgauw (NL).

(74) Agents: DE BRUIJN, Leendert, C. et al.; Nederlandsch Octrooibureau, Scheveningseweg 82, P.O. Box 29720, NL-2502 LS The Hague (NL).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: METHOD FOR THE CONSTRUCTION OF VECTORS FOR LACTIC ACID BACTERIA LIKE LACTOBACILLUS SUCH THAT THE BACTERIA CAN EFFICIENTLY EXPRESS, SECRETE AND DISPLAY PROTEINS AT THE SURFACE



(57) Abstract

The invention relates to a method to construct multi-purpose plasmid vectors that can be used for the introduction, stable maintenance and efficient expression and secretion of foreign genes or gene fragments in a variety of lactic acid strains such as *Lactobacillus* strains and *Bifidobacterium*, as well as to expression vectors to be used in this method. The invention also comprises a method of production of proteins and polypeptides heterologous and homologous to lactic acid bacteria at increased levels. Also comprised are lactic acid bacteria containing these expression vectors and producing these proteins or polypeptides.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

METHOD FOR THE CONSTRUCTION OF VECTORS FOR LACTIC ACID BACTERIA LIKE LACTOBACILLUS SUCH THAT THE BACTERIA CAN EFFICIENTLY EXPRESS, SECRETE AND DISPLAY PROTEINS AT THE SURFACE

FIELD OF THE INVENTION

The present invention relates to a method to construct multi-purpose plasmid vectors that can be used for the introduction, stable maintenance and efficient expression and secretion of foreign genes in a variety of lactic acid strains such as *Lactobacillus* strains and *Bifidobacteria*, as well as to expression vectors to be used in this method. The invention also comprises a method of production of proteins and polypeptides heterologous to lactic acid bacteria at increased levels as well as some novel products.

BACKGROUND OF THE INVENTION

Introduction

Recombinant DNA technology has opened the way to an efficient production of proteins and enzymes from human, animal, plant or bacterial sources in a variety of organisms. During the past fifteen years it has become routine to (over)produce a protein of interest by cloning the corresponding gene onto a (multi-copy) plasmid vector, and expressing the gene under the control of strong expression signals, in a microorganism such as for example the bacterium *Escherichia coli*. Application of this methodology to other bacterial hosts, in particular industrially important bacteria like Lactic Acid Bacteria (LAB), more specifically *Lactobacilli*, requires that cloned DNA sequences can be stably maintained and expressed. Over the past decade systems have been developed with which homologous and heterologous genes can be introduced in *Lactobacilli* on (multi-copy) plasmid vectors. Such hybrid vectors are, in general, structurally stable and can be stably maintained under selective conditions. Systems for site-specific integration of cloned genes have also been described [Chopin et al, Appl. Environm. Microbiol. 55, 1769-1774 1989; Leenhout et al, Appl. Environm. Microbiol. 55, 394-400 1989; Leer et al, Mol. Gen. Genet. 239 269-272 1993]. Cloned genes that are integrated into the *Lactobacillus* chromosome are both structurally and segregationally stable, in the presence or absence of selective pressure.

Heterologous gene expression in *Lactobacillus*

Rapid advances in applying gene technology for the development of strain improvement programmes for *Lactobacilli* are limited by i) structural instability of *Lactobacillus* DNA in working organisms like *Escherichia coli* and ii) inefficiency of expression and secretion of heterologous proteins in *Lactobacillus*.

Structural instability of *Lactobacillus* DNA sequences cloned in *Escherichia coli* was reported in a number of publications [Bates and Gilbert, Gene 85, 253-289 (1989); Taguchi and Ohta, J. Biol. Chem. 266, 12588-12594 (1991); Bernard et al, FEBS Lett. 290, 61-64 (1991); Kochhar et al, Biochem. Biophys. Res. Commun. 185, 705-712 (1992); Lee et al, J. Biol. Chem. 266, 13028-13034 (1989); Lerch et al, Gene 78, 47-57 (1989); Vanderslice et al, J. Biol. Chem. 261, 15186-15191 (1986); Posno et al, Appl. Environm. Microbiol. 57, 2764-2769 (1991); Branny et al, J. Bacteriol. 175, 1993]. In general, instability increases when cloned *Lactobacillus*-derived DNA fragments are expressed in *Escherichia coli*. For example when the *xylA* gene of *Lactobacillus pentosus*, encoding xylose isomerase, was introduced in *Escherichia coli* on a high copy vector (pUC19) in two parts (one of which also contained the promoter region), these recombinant vectors were structurally stable. However, reconstitution in *Escherichia coli* of the entire *xylA* gene from the two parts proved to be impossible [Posno et al, Appl. Environm. Microbiol. 57, 2764-2766 (1991)]. Similarly, reconstitution of the *Lactobacillus plantarum* D-lactate dehydrogenase gene by joining parts of the gene present on two separate clones could not be achieved in *Escherichia coli*. It was suggested that the product of the intact D-*ldh* gene was toxic for *Escherichia coli* [Taguchi et al, J. Biol. Chem. 266, 12588-12594 (1991)]. This type of instability can be explained in general by assuming that (over)production of a *Lactobacillus* gene product is toxic for *Escherichia coli*.

Structural instability of plasmid vectors with strong cloned promoters has been reported by different groups [e.g. Gibson et al, Gene 53, 275-281 1987] and has also been observed for the promoter of the *Lactobacillus casei ldh* gene by the present inventors. This type of instability is explained by assuming that transcription starting at the promoter interferes with efficient replication of the plasmid vector. The instability can, in general, be overcome by insertion of a transcription

terminator sequence between the promoter and the origin of replication.

Although the inconvenience of structural instability during construction of recombinant DNA molecules and performance of other necessary genetic manipulations might be overcome by using another host organism, e.g. *Lactobacillus* itself, the bacterium *Escherichia coli* is preferred as "working organism" because of the wealth of genetic and biochemical data that are available concerning this organism. In addition, genetic manipulation of lactic acid bacteria like *Lactobacilli* is characterized by intrinsic difficulties which make this host less suitable for standard recombinant DNA experiments than a standard working organism like *Escherichia coli*.

European Patent Application EP-B-0.449.770 describes a shuttle vector directed at secretion of homologous and heterologous polypeptide or protein from Lactococcus or Bacillus microorganisms. Secretion is deemed necessary due to inherent problems of internal storage of heterologous expression products requiring refolding of the expression product. Nothing is stated or suggested with regard to instability of a lactic acid bacterial expression product or nucleic acid sequence in a working organism such as E. coli. Nothing is stated or suggested with regard to expression problems of heterologous genes in Lactobacilli or Bifidobacteria. Nothing is stated or suggested regarding amount of expression product produced by a lactic acid bacterium host cell. The disclosure of the cited patent application is directed at the use of the signal peptide of the Lactococcus MSP (=Major Secretion Product) for obtaining secretion of the expression product as opposed to storage intracellularly. Nothing is taught or suggested regarding the required presence of any lactic acid bacterial nucleic acid sequences for obtaining high amounts of expression product. The problem adressed in the cited patent application differs from the problem addressed in the instant case. Vectors described in the abovementioned European Patent Application do not fall within the category claimed in the instant case.

In WO 93/17117 heterologous polypeptides are produced in Lactococcus by means of a T7 or T7-like RNA polymerase gene placed under the control of an inducible promoter effective in the Lactococcal host and a promoter specific for said polymerase upstream of a coding sequence for the desired polypeptide, whereby the promoter directs transcription of said

coding sequence selectively as a result of expression of said polymerase. The T7 RNA polymerase system is an E coli system that also works in Lactococcus. Nothing is mentioned or suggested concerning expression of heterologous protein or polypeptide in other lactic acid bacteria such as Lactobacilli and Bifidobacteria. Nothing is taught or suggested regarding the required presence of any lactic acid bacterial nucleic acid sequences for obtaining high amounts of expression product. The only details regarding the presence of lactic acid bacterial sequences relate to the presence of lactococcal signal peptide for secretion of the expression product by Lactococcus. This cited document teaches that a high amount of expression product is attained due to the use of the powerful T7 promoter which is transcribed at a high level in the presence of the specific polymerase so that transcription is not substantially effected by other polymerases in the cell and can be controlled by controlling the inducible promoter thus controlling the expression of the specific polymerase. The cited patent application does not address the issue of incompatibility of lactic acid bacterial expression products or nucleic acid sequences in a working organism such as E. coli. More specifically a promoter operable in E coli, the T7 promoter is described linked to the signal peptide and protein translation initiation signal of the lactococcal protease gene, whereby the presence of the protease sequence is to ensure secretion of the expression product. The subject invention however is not directed at expression vectors that are only useful in Lactococcus or E coli. Neither is the subject invention directed at improved expression through use of a specific strong promoter.

The cited patent application provides a good summary of the state of the art regarding lactic acid bacterial expression of heterologous proteins or polypeptides and clearly illustrates that the problem of low expression using lactic acid bacteria as expression host cells has existed for a long time and that the state of the art has not taught or suggested a solution along the lines indicated by the subject invention which offers a general solution to the problem. Use of the T7 promoter as described in the cited state of the art is not comprised within the subject invention.

Preliminary studies which deal with expression of heterologous proteins in lactic acid bacteria like *Lactobacilli* have shown that expression of foreign proteins, in particular mammalian proteins, in *Lactobacillus* is

rather inefficient. In none of the studies published until WO93/17117 was expression of heterologous nucleic acid sufficiently efficient to obtain levels of production that surmounted 1-2% of total soluble protein when using lactic acid bacteria. The subject inventors have observed in particular that the efficiency of expression in *Lactobacillus* e.g. of the *Escherichia coli* enzyme β-galactosidase was in fact reduced ten- to twenty-fold (final level 0.1-0.2% of total soluble protein) when β-galactosidase was fused at its N-terminal end with a twenty to thirty amino acid antigenic determinant from the Foot and Mouse Disease Virus capsid protein VP1 or the Rotavirus capsid protein VP7.

In WO94/00581 an expression and secretion system of recombinant genes in Gram positive bacteria and especially Lactobacillus is described. This cited patent application clearly states that low level heterologous expression of α-amylase, ß-glucanase and endoglucanase genes utilizing their endogenous expression units and TEM-ß-lactamase utilising a lactococcal promoter had been reported in the state of the art. in addition it is disclosed that no reports of the characterisation of Lactobacillus expression units or the use of such expression units for homologous or heterologous expression in Lactobacillus were known. The cited application subsequently describes the characterisation and utilisation of the export signals of SP (=surface layer protein) of Lactobacillus brevis for export or secretion of homologous or heterologous gene products. It is stated that the selection of this protein to provide efficient regulatory elements was due to the large amounts in which the protein was known to be present in Lactobacillus and its presence on the outside of the host membrane along with the supposition it must be a constitutive promoter. It also states that the promoter unit of SP can be used alone devoid of the ribosome binding site of the SP gene. No illustration of amounts of expression products using the SP regulatory units is given. The cited patent application does not address the issue of incompatibility of lactic acid bacterial expression products or a nucleic acid sequence in a working organism such as E. coli. This cited patent application clearly illustrates that the problem of low expression using lactic acid bacteria as expression host cells has existed for a long time and that the state of the art has not taught or suggested a solution along the lines indicated by the subject invention which offers a general solution to the problem. Use of the SP regulatory elements as described in the cited state of the art is not comprised

within the subject invention.

In WO 92/04451 and US 5,242,821 a number of specific Lactococcus promoters and promoter/secretion promoting signals are disclosed as being useful in production of heterologous and homologous proteins and peptides in E. coli and especially in Gram positive bacteria. The promoter sequences are those found in plasmids pKTH1789, pKTH1816, pKTH1817, pKTH1820, pKTH1821 and pKTH1874. The promoter and secretion promoting signals are those found in plasmids pKTH1797, pKTH1798, pKTH1799, pKTH1801, pKTH1805 and pKTH1806, pKTH1807 and pKTH1809. The promoter/signal promoting sequence is described as being "generally a sequence encoding a 16-35 amino acid segment usually containing hydrophobic amino acids that become embedded in the lipid bilayer membrane which allows for the secretion of an accompanying protein or peptide sequence". The regulatory elements of these plasmids can according to this document be recombined to produce hybrid expression units which can function together to allow enhanced heterologous expression in E. coli and Grampositive bacteria. The description further illustrates how promoters can be picked up by a probe vector lacking a promoter prior to a cat gene and how secretion signal sequences can be picked up using a vector comprising the marker gene ß-lactamase. No indication is given of other heterologous or homologous proteins or polypeptides to be expressed. No comparison of degree of expression with existing systems is given. No illustration with other sequences than Lactococcus sequences is given. No indication is given for different degrees of expression from the various promoters. The invention provides a method for randomly screening for regulatory sequences of Lactococcus. Once such sequences have been ascertained and isolated they can be introduced and applied in other vector systems. No expression vector comprising a lactic acid bacterium promoter followed by at least the first five codons of a lactic acid bacterium gene followed by a further gene product under regulation of the same promoter is disclosed explicitly as such for improving expression of the further gene in lactic acid bacteria in the instant cited application. However a vector being any one of plasmids pKTH1797, pKTH1798, pKTH1799, pKTH1801, pKTH1805, pKTH1806, pKTH1807, pKTH1809 and pKTH1889 does comprise a promoter and a signal sequence from lactococcus regulating the ß-lactamase gene. The signal sequence comprises the first codons of the relevant lactococcal gene from which it is derived for secretion purposes of ß-lactamase. Such

vectors are however not used for increased expression of heterologous genes in lactic acid bacteria but as probe vectors. Whether the intermediate lactococcal signal sequence could in fact be responsible for increased expression of a heterologous gene linked thereto is also not disclosed. Neither is any indication given of the nature of the intermediate sequence other than that it being a lactococcal secretion signal as described above. In fact it is clearly stated for vectors with promoters of the selected vectors in combination with the signal sequence when judged by transcriptional efficiency that "the promoters cloned by the cat plasmid pKTH1750 appeared stronger than promoters cloned together with the signal sequence the test signal gene being bla in the latter case". The state of the art vectors pKTH1797, pKTH1798, pKTH1799, pKTH1801, pKTH1805, pKTH1806, pKTH1807, pKTH1809 and pKTH1889 are not included within the scope of protection of the subject vector claim.

Object of the invention

The objective of the present invention is to overcome the above mentioned difficulties and to provide per se and in combination

i) a method for the construction of plasmid expression and expression-secretion vectors in a working organism other than the lactic acid bacterium to be transformed, for example in *the working organism Escherichia coli*,

ii) a series of expression and expression-secretion vectors for transforming lactic acid bacteria like *Lactobacillus* and Bifidobacteria, said vectors being characterized by at least a translation initiation region allowing efficient expression of heterologous nucleic acid encoding foreign proteins or polypeptides, and

iii) a series of vectors that efficiently express, secrete and display antigens from pathogens such as influenza virus or rotavirus at the bacterial surface of their host, said vectors preferably being food grade. Thus providing a method for expressing and optionally secreting and displaying heterologous proteins or polypeptides in lactic acid bacteria strains. In particular expression of heterologous proteins or polypeptides in Lactobacilli and Bifidobacteria and most preferably in Lactobacilli is an objective of the subject invention. The products thus obtainable are also considered to fall within the scope of the invention.

SUMMARY OF THE INVENTION

The present invention provides a method to construct by recombination DNA

technology a series of expression vectors and expression-secretion vectors in a working organism other than the lactic acid bacterium to be transformed, in particular a non *Lactobacillus* working microorganism such as *Escherichia coli* that can be used to genetically modify a bacterium, in particular a lactic acid bacterium so as to efficiently produce foreign proteins or polypeptides. Especially interesting microorganisms for expression are Lactobacilli and Bifidobacteria. Especially interesting is the production of immunogenic proteins and peptides from pathogenic organisms and proteins e.g. mammalian proteins that can be used for the induction of immunological tolerance. Other interesting objectives are production of micronutrients, vitamins and prebiotic factors. In particular products that could now be obtained are human and pig rotavirus envelope proteins, influenza virus haemagglutinins, influenza virus nucleo proteins, heat labile toxin of *E. coli* and fusion proteins thereof, cholera toxin (ß-subunits) and fusion proteins thereof, urease A and/or B from *Helicobacter pylori*. The nucleic acid sequences encoding these proteins form part of the state of the art and recombinant lactic acid bacterial expression of such sequences will lead to products that are illustrative of the products that have now become available. In particular the fact that such products can be obtained as more than 1%, even more than 2% and in particular more than 3% of the total soluble protein content of a lactic acid bacterium like *Lactobacillus* in a simple manner using an expression vector according to the invention is particularly interesting. It now also becomes possible to provide oral vaccines and immunologically stimulating preparations. Use of a foodgrade host cell and food grade expression vector enables production of food grade proteins and polypeptides in particular of proteins and polypeptides that are heterologous to the lactic acid bacterium host cell or even to lactic acid bacteria per se.

It was found that the level of expression of a heterologous protein or polypeptide in a lactic acid bacterium, whilst the corresponding nucleic acid is under control of a promoter operative in the lactic acid bacterium, can be considerably increased by insertion of a nucleotide sequence coding for the N-terminal sequence of a gene homologous to a lactic acid bacterium gene, immediately downstream of the translation start codon. The intermediate lactic acid bacterium sequence need only be 5 codons long but can be anywhere in length between 5-15 codons, even 5-34 codons. The fragment can even comprise up to the 5' half of the gene.

It will be preferred in general to keep the fragment as short as possible. A complete gene without transcription termination could in theory also be applied as intermediate lactic acid bacterium sequence, but in practice does not seem overly attractive. The protein or polypeptide can be homologous or heterologous to the lactic acid bacterium in which it is to be expressed, with a preference for heterologous. The 5' non translated region located between the intermediate sequence and the promoter is preferably homologous to lactic acid bacterium, in particular preferably homologous to the host cell that is to comprise the expression vector for expression of the protein or polypeptide that is to be expressed jointly with the 5' intermediate lactic acid bacterium sequence. It is efficient for the 5' non translated and 5' translated lactic acid bacterium sequences to be derived from the same organism and indeed preferred for them to be derived from the same gene. The same is valid for the promoter regulating the expression of the intermediate sequence and the protein and polypeptide to be expressed in the vector according to the invention. Suitably promoter, 5' non translated region and 5' terminal translated fragment of at least 5 codons are all derived from the same lactic acid bacterium and preferably the same gene. The latter saves the person skilled in the art finding the operable sizes and distances of the relevant components for good expression. Naturally however the sizes and distances between the various components can be varied in an attempt to optimise expression. The basic order of the components is however not subject to alteration.

The present invention more specifically provides:

i) A series of plasmid vectors with which foreign proteins, in particular immunogenic proteins from pathogenic organisms or mammalian proteins that can induce immunological tolerance, can be efficiently expressed in *Lactobacillus*. These foreign proteins can optionally be accumulated or excreted and/or secreted either into the culture fluid or subsequently anchored to the membrane and displayed at the surface of bacteria. The expression in particular in and on *Lactobacilli* of such heterologous expression products is of interest,

ii) A method to construct plasmid vectors as described in i) in a working organism other than the type in which the heterologous product is to be expressed, whereby structural instability of the *lactic acid bacterial vectors, more specifically of the Lactobacillus* vectors is circumvented. Such a working organism can suitably be *Escherichia coli*.

iii) Lactic acid bacteria strains harbouring plasmid vectors that efficiently express antigenic determinants from e.g. influenza virus and rotavirus fused to an immunological carrier protein like β-glucuronidase. Specifically of interest are *Lactobacillus* strains.

DETAILED DESCRIPTION OF THE INVENTION

i) Facilitated construction in a non lactic acid bacterium for example in a working organism like *Escherichia coli*

The present invention is characterized by a novel method to construct the vectors in a non lactic acid bacterium for example in a working organism like *Escherichia coli* prior to their introduction in a lactic acid bacterium like *Lactobacillus*. Application of the invention thus in particular circumvents the structural instability of *Lactobacillus* vector material in *Escherichia coli* caused by cloning of strong *Lactobacillus* promoters and/or by cloning and expression in *Escherichia coli* of *Lactobacillus* genes. The construction route in a non lactic acid bacterium for example in a working organism like *Escherichia coli* involves the insertion of a nucleotide sequence containing a transcription terminator between the promoter controlling the expression of the heterologous nucleic acid and the transcription terminator for the heterologous gene located downstream from the promoter. The newly introduced terminator serves to prevent transcription in *Escherichia coli* of open reading frames (ORF) that are to be inserted downstream from the promoter sequence so as to prevent their expression in *Escherichia coli*. Preferably the newly introduced terminator sequence will be located such that no product controlled by the promoter is produced in the working organism. This can be achieved by locating the newly introduced terminator prior to the first codon of the heterologous expression product. A person skilled in the art will realise that in some instances a short expression product, i.e. a partial amino acid sequence expressed from lactic acid bacterium nucleic acid will be tolerated by the working organism, so that location of the newly introduced terminator sequence within the heterologous sequence encoding heterologous protein or polypeptide could be acceptable to prevent the negative consequences that expression of the complete product would otherwise have on the working organism. The newly introduced terminator sequence is preferably a strong terminator sequence. The newly introduced terminator sequence must be recognised as such by the working organism in which the vector is to be constructed or manipulated prior to transformation of a lactic acid

bacterium. A preferred embodiment comprises a newly introduced terminator sequence that is recognised by the working organism, a non lactic acid bacterium but is not recognised by the lactic acid bacterium in which expression of the heterologous protein or polypeptide is desired.

In particular when a newly introduced terminator sequence is employed that can also be recognised by a lactic acid bacterium in which the heterologous protein or polypeptide is to be expressed it is necessary for the transcription block to be lifted. In order to achieve this the newly introduced terminator sequence is flanked at both sides such that it is excisable from the vector in a manner known per se to a person skilled in the art. This can be the case for example when it is flanked by a nucleic acid sequence encoding one or more very infrequently occurring restriction enzyme sites (e.g. *Not*I) which do not occur elsewhere in the plasmid, allowing easy removal of the newly introduced terminator sequence from the vector prior to its introduction into the lactic acid bacterium for example a *Lactobacillus*. Alternatively the newly introduced terminator sequence may be flanked by direct repeats such that the terminator sequence is excisable through homologous recombination of the host. After removal of the newly introduced terminator sequence the cloned ORF with the heterologous protein or polypeptide will be transcribed from the promoter in the lactic acid bacterium.

A person skilled in the art will realise that the working organism need not necessarily be a non lactic acid bacterium; this is merely a preferred embodiment due to the inherent problems of working in a lactic acid bacterium as working organism. Furthermore a person skilled in the art will realise that the nucleic acid sequence encoding a protein or polypeptide i.e. the product to be expressed need not necessarily be heterologous to a lactic acid bacterium. The system according to the invention may also be employed for homologous nucleic acid sequences. Generally speaking however expression of heterologous sequences normally evokes more problems than expression of homologous sequences.

ii) Enhanced expression in *Lactobacillus*

The present invention also provides expression and expression-secretion vectors that allow efficient expression of cloned foreign genes. It was found that the level of expression of a heterologous protein or

polypeptide like the *Escherichia coli* enzyme β-galactosidase in a lactic acid bacterium such as *Lactobacillus*, whilst the corresponding nucleic acid is under control of a homologous promoter of the lactic acid bacterium, for example a *Lactobacillus* promoter, can be considerably increased by insertion of a nucleotide sequence coding for the N-terminal sequence of a gene homologous to a lactic acid bacterium like a *Lactobacillus* immediately downstream of the translation start codon.

In particular it is postulated that the presence of A and T rich nucleic acid sequences at the beginning of a nucleic acid sequence to be expressed in a lactic acid bacterium is conducive to high expression being achieved from such sequences. As in particular lactic acid bacteria like *Lactobacillus* have A and T rich sequences in parts of their genes encoding the N terminal part of their proteins this could be the reason why the presence of such a sequence ensures good expression in a lactic acid bacterium. It is possible that a high number of G and C residues in this region would favour the formation of stable RNA duplexes, which may interfere with translation initiation. Preferred fragments located immediately behind the initiation codon and prior to the protein or polypeptide to be expressed have an AT percentage of 60% or more, preferably higher than 62% , also 65% or more and even 70% or more. Even higher percentages can be used. AT percentages as high as 85 and 90% even up to 95% and as high as 100% are suitable. As discussed above the length of the fragment is not critical, but must be at least 5 codons. It is preferable to use sequences that naturally occur in a lactic acid bacterium, preferably that occur in the lactic acid bacterium host cell in which the protein or polypeptide has to be expressed. The Lactobacillus microorganism has to date been a most difficult organism in which to achieve good stable expression in particular of nucleic acid heterologous to Lactobacillus or to any lactic acid bacterium and as such such an organism is a preferred host cell for expression vectors according to the invention.

In order to obtain good expression in a lactic acid bacterium and optionally surface display of a protein or polypeptide, said vector must comprise an expression promoter sequence and a nucleic acid sequence encoding said protein or polypeptide with expression of said nucleic acid sequence being controlled by said promoter sequence, characterised in that said encoding nucleic acid sequence is preceded by a 5' non-

translated nucleic acid sequence comprising at least the minimal sequence required for ribosome recognition and RNA stabilisation followed by a translation initiation codon, said translation initiation codon being immediately followed by a fragment of at least 5 codons of the 5' terminal part of the translated nucleic acid sequence of a gene of a lactic acid bacterium or a structural or functional equivalent of said fragment, said fragment also being controlled by said promoter, thus resulting in an expression product which is a fusion polypeptide. As described in the introduction some vectors or plasmids have inadvertently been described in the state of the art comprising the elements minimally required to achieve the desired result of the invention. These known vectors have been excluded from the scope of protection of the subject claims.

A structural or functional equivalent of said fragment of at least 5 codons is understood to include fragments encoding the same amino acid sequence but using different codons, fragments derived from other organisms than lactic acid bacteria but encoding corresponding parts of functionally related polypeptides, and equivalents of said fragment having at least 60% or even 80% nucleotide homology and preferably containing at least all the A and T residues of the naturally occurring fragment are covered by this definition. Preferablt the native sequences will be applied.

The gene from which the 5' terminal part of the translated nucleic acid sequence is derived is preferably the same gene from which the 5' non-translated region located between the promoter which is used to control the expression of the heterologous nucleic acid is derived. The resulting expression product thus becomes a hybrid of homologous and heterologous amino acid sequences, whereby the presence of the N-terminal sequence of the lactic acid bacterium gene in the resulting expression product does not detract from the functionality of the heterologous product that needs to be expressed. It is not necessary for the homologous N-terminal sequence to be very long. A nucleic acid sequence encoding at least 5 amino acids is sufficient to obtain good expression results. The maximum length to be used will depend on the expression host and the nature of the expression product. The maximum length is not really critical. One embodiment can comprise as small a sequence as possible in order not to risk the loss of functionality of the resulting expression product due to

tertiary folding like e.g. proteolytic instability sufficiently to disturb the activity of the heterologous protein or polypeptide. A sequence of ten amino acids has been found to form an embodiment that is capable of achieving the objective of the invention quite adequately. N-terminal sequences with lengths over 20 amino acids are acceptable embodiments as well. Alternatively longer sequences are contemplated embodiments forming e.g. complete genes, without the corresponding gene terminator. By way of example even polycistronic mRNA encoding at least two proteins can be envisaged.

The expression and expression-secretion vectors of the present invention can comprise the entire non-translated 5' region including the Shine-Dalgarno sequence [Shine and Dalgarno, Proc. Natl. Acad. Sci. USA 71, 5463-5467 (1974)], translation initiation codon and the A/T-rich N-terminal sequences preferably of a lactic acid bacterium gene. The selection of the gene from which the N-terminal sequence and the promoter are to be used will be obvious to a person skilled in the art. The criteria that are relevant are the organism in which expression is desired and the degree of expression that is normal for the gene system of which some components are to be used in the vector when the gene system is functioning normally in its natural environment. For *Lactobacillus* for example a number of proteins are known or are expected to be expressed to a high degree and any of the relevant components from the corresponding gene systems of *Lactobacillus* are suitable for use in vectors according to the invention. A large deal of sequencing has already taken place and a person skilled in the art can quite readily ascertain suitable sequences that can be used. The lactate dehydrogenase encoding gene, *ldh*, the proteinase encoding gene, *prtP*, or the α-amylase encoding gene, *amy* the S layer protein encoding gene slpA (Boot et al. Journ. of Bacteriol. Oct. 93 p. 6089-6096) (Figure 12), the ribosomal RNA encoding gene rrn, the glyceraldehyde-3-phosphate dehydrogenase encoding gene gapdh are examples of genes that are efficiently expressed in *Lactobacillus*. Their promoters and N terminal sequences can quite suitably be employed in vectors according to the invention. The configuration around the translation initiation start site in the expression vectors is chosen so as to warrant translation of foreign proteins in a manner equally efficient to that of the authentic *Lactobacillus*. L-lactate dehydrogenase, proteinase, α-amylase, S layer protein, the ribosomal RNA and the glyceraldehyde-3-phosphate

dehydrogenase are all proteins that are efficiently translated in *Lactobacillus*. L-Lactate dehydrogenase is a key enzyme in the glycolytic pathway of lactobacilli which mediates the conversion of pyruvate to lactate. The proteinase encoded by *prtP* is an enzyme secreted and anchored to the cell surface through a so-called anchor sequence by a number of *Lactococci* and *Lactobacilli* which enables the bacteria to proteolytically hydrolyse casein and therefore grow on milk. Finally, α-amylase is an enzyme secreted by *Lactobacillus amylovorus* that enables the bacterium to grow on starch as sole energy source.

Furthermore a person skilled in the art will realise that the nucleic acid sequence encoding a protein or polypeptide i.e. the product to be expressed need not necessarily be heterologous to a lactic acid bacterium. Naturally the system according to the invention may also be employed for expression of homologous nucleic acid sequences.

iii) Preferred embodiments

-a- cassette structure

The present invention is further in preferred embodiments characterized by the use of fully sequenced, multi-copy plasmid expression and expression-secretion vectors carrying a replicon from the lactic acid bacterium in which the heterologous nucleic acid is to be expressed. Such an organism can be a *Lactobacillus*. The vectors preferably have a modular structure facilitating easy manipulation and exchange of cassettes from one vector to another. The cassettes can comprise in addition to the essential components described in the preceding description of embodiments of the invention and those components of neccesity present in a vector in order to be operable a number of additional components. The vector includes in the order 5' to 3' one or more additional components. the additional components follow the constitutive promoter or the strong regulatable promoter, which promoter could e.g. be from a lactic acid bacterium such as *Lactobacillus* and/or follow the transcription termination sequence for blocking transcription of the heterologous protein or polypeptide in the working host. The components can be

i) a DNA sequence encoding the signal sequence of a secreted protein recognisable by a lactic acid bacterium as such, this could for example be a signal sequence from a lactic acid bacterium,

ii) a reporter gene encoding an easily assayable product in the working organism like e.g. β-glucuronidase from *Escherichia coli*, said reporter

gene can be located either in frame between the initiation codon and the nucleic acid sequence encoding a protein or a polypeptide or in frame after the nucleic acid sequence encoding a protein or a polypeptide under the control of the same promoter as the nucleic acid sequence encoding a protein or polypeptide,

iii) a DNA sequence encoding an anchor sequence from a gene of a lactic acid bacterium for example like *Lactobacillus*, said DNA sequence being located in frame downstream from the nucleic acid sequence encoding a protein or a polypeptide and

iv) a transcription termination sequence of a lactic acid bacterium gene like a *Lactobacillus* gene located such as to ensure only expression of the desired heterologous protein or polypeptide and

v) A DNA sequence encoding a probiotic product.

The working host is preferably a non lactic acid bacterium, in particular not a *Lactobacillus* and preferably an E coli.

-b- selected cloning sites

Additional embodiments of the invention are further characterized by the presence in the vectors, between the translation initiation region and reporter gene, of a series of restriction enzyme sites (MCS) in three reading frames, allowing in-phase translation of any foreign DNA sequence, coding for e.g. an antigenic determinant. The advantages of MCS in vectors are well known to a person skilled in the art. In the distal part of the MCS the present restriction enzyme sites can be easily manipulated to allow in-phase translation of inserted ORF sequences with the downstream reporter gene; or when the reporter gene is omitted, in-phase fusion with an anchor sequence if immobilisation on the surface of the organism in which the sequence is expressed is required. Such an anchor sequence can be that of prtP of *Lactobacillus* if *Lactobacillus* is to be transformed. In phase fusion may also be desirable when the product to be expressed should not contain additional amino acids at the carboxyl end. This can be realised through fusion to the last amino acid codon of the *cbh* gene.

These restriction enzyme sites can be easily manipulated so as to allow in-phase translation of the inserted DNA sequence with a downstream reporter gene. A further embodiment of the invention is the presence in the vector of restriction enzyme sites that permit easy deletion of the

reporter gene. Another option is the presence of sites such that substitution of the reporter gene by other genes coding for desired proteins, polypeptides or parts thereof can occur. Such proteins or polypeptide can serve the purpose of an immunogenic carrier, e.g. *Escherichia coli* β-galactosidase can be suitably used for this purpose.

-c- reporter gene

Further embodiments of the present invention are additionally characterized by the possibility to fuse the antigenic determinant of interest to a large protein moiety (e.g. β-glucuronidase) which serves two purposes. Because of its ability to hydrolyse the chromogenic substrate X-gluc (5-bromo-4-chloro-3-indolyl-β-D glucuronic acid), β-glucuronidase can serve as reporter protein. Synthesis of antigenic determinants fused with β-glucuronidase can easily and semiquantitatively be monitored after growth of transformed bacteria on solid media, supplemented with X-gluc. Beta-glucuronidase activity can be quantitatively determined after growth of the bacteria in liquid media. Fusion of an antigenic determinant to a large immunogenic protein (e.g β-glucuronidase), enhances the immune response against the antigenic determinant, in particular if the determinant itself is small. The immunogenic protein serves as the carrier. The cassette structure allows deletion of the reporter gene at will, for example for expression of non-fused heterologous proteins.

Another option is to include nucleic acid encoding products with health stimulating properties to further enhance the value of a transformed lactic acid bacterium. Examples of such proteins are immunomodulators, e.g. cytokines, bacteriocines, glutathione-S-transferase, immuno-globulins and levanase. The sequences encoding such proteins are comprised in the state of the art or can be derived without undue burden in a number of cases from clones comprising the relevant sequences. The sequences of such products are hereby enclosed by reference.

-d- site of protein expression

Embodiments of the present invention can additionally be characterized by the possibility, in the case of the synthesis of proteins that are to be secreted and surface-bound e.g. via the membrane or the S layer protein. For binding to the membrane the length of the spacer region between the antigenic determinant and the anchor sequence can be varied, depending on

the thickness of the peptidoglycan layer and the sensitivity of the fusion protein to proteolytic attack.

-e- food-grade expression vector

Embodiments of the present invention can further be characterized by the possibility to replace commonly used antibiotic resistance markers and *Escherichia coli* DNA sequences of the vectors by a food grade marker e.g. a nutritional selection marker from a food-grade organism, to obtain a food-grade vector comprising elements derived from food-grade organisms only. The food grade marker preferably originates from *Lactobacillus*, enabling the construction of vectors to be performed under conditions that are exempt from regulations for recombinant DNA research of the EC.

Basic vector construction for a vector suitable for Lactobacillus

Firstly as starting material for the construction of the expression and expression-secretion vectors, pGEM-3 was used. The strategy for the construction of the expression vector is as follows. Expression from strong *Lactobacillus* promoters like the *Lactobacillus casei ldh* promoter in *E. coli* frequently leads to structural instability which can be overcome by inserting a terminator directly behind the promoter sequence. Consequently, the transcription terminator (e.g. T_{cbh}) was first inserted in pGEM-3 before a *Lactobacillus* promoter sequence was introduced. In addition, as expression of genes under the control of *Lactobacillus* promoters in *Escherichia coli* frequently results in structural instability, the transcription terminator sequence of the *Lactobacillus casei ldh* gene was inserted between the promoter and the *Lactobacillus plantarum cbh* terminator, prior to the introduction of the *Escherichia coli uidA* gene.

The nucleotide sequence of pGEM-3 (Promega), of the *Lactobacillus* plasmid vector pLPE323 [Leer et al, Mol. Gen. Genet. 234, 265-274 1992] and of all other elements used for the construction of the expression vectors, such as promoter sequences, terminator sequences etc, is fully known. All elements were provided with flanking nucleotide sequences harbouring restriction enzyme sites which do not occur elsewhere in the plasmid, permitting easy manipulation of the elements.

Multi-purpose cassette construction

The final set of expression cassettes was equipped with a number of

unique restriction enzyme sites permitting further modification of the vectors, if wanted. To achieve this, a 260-bp *Hpa*I-*Hin*dIII fragment containing the last amino acid codon and transcription terminator sequence of the *cbh* gene of *Lactobacillus plantarum* 80 [Christiaens et al, Appl. Environm. Microbiol. 58, 3792-3798 1992] was ligated between the *Hin*cII and *Hin*dIII sites of pGEM-3. Subsequently, the *Hin*dIII site was converted into an *Eag*I site using a synthetic linker sequence. The remaining multiple cloning sequence was replaced by a chemically synthesized oligonucleotide (PRESS) containing the following restriction enzyme sites, *Eco*RI, *Bgl*II, *Spe*I, *Sph*I, *Bam*HI, *Bst*EII. The *Bst*EII site is adjacent to the *Hin*cII site used for cloning of the *cbh* fragment. The resulting vector which was designated pGTC3 (Figure 1) was used for the insertion of three DNA fragments containing expression and expression-secretion sequences.

Secondly, promoter bearing DNA fragments were inserted into the new PRESS of pGTC3 (Figure 1).

i) The promoter, translation initiation region and first 8 specific codons of *Lactobacillus casei ldh*, which are present on a 513-bp *Sau*3AI fragment were cloned between the *Bgl*II and *Bam*HI sites of pGTC3, leaving the *Bam*HI site 3' from the promoter sequence intact. First, a DNA fragment containing the 5' part of the *Lactobacillus casei ldh* gene [Kim et al, Appl. Environm. Microbiol. 57, 2413-2417 1991] was synthesized by PCR and used as a probe to isolate a 3.4-kb *Pst*I-*Eco*RI fragment from a partial genomic library cloned into pUC19. This fragment was used to generate the 513-bp *Sau*3AI fragment (Figure 1). The resulting vector comprising pGTC3 and the expression sequence of the L-*ldh* gene of *Lactobacillus casei* ATCC 393 was called pGLC3 (Figure 2A-b).

ii) To obtain a P_{ldh}-directed expression system with which an expressed protein can also be secreted, the secretion signal sequences of the *Lactobacillus casei* ATCC 393 *prtP* gene were isolated by PCR. With primers based on the *ptrP* gene of *L. lactis* [Kok et al, Appl. Environm. Microbiol. 54, 231-238 1988] and *Lactobacillus paracasei* subsp. *paracasei* NCDO 151 [Holck and Naes, J. Gen. Microbiol.138, 1353-1364 1992] a PCR fragment from *Lactobacillus casei* ATCC 393 was amplified which contains the Shine-Dalgarno sequence, translation startcodon and the first 70 codons of the *prtP* gene. The identity of this fragment was confirmed by DNA sequence analysis of the cloned fragment. From this cloned PCR

fragment a 239-bp *Sph*I-*Bam*HI fragment was isolated and used to replace the 46-bp *Sph*I-*Bam*HI fragment containing the translation initiation region and first 8 codons of the *ldh* gene. The resulting vector was named pGLC3-ss (Figure 2A-c).

iii) A *Bgl*II-*Sph*I fragment (864 bp) from pLPCR2-1 [Jore et al, 1995 submitted for publication] containing the transcription initiation sequences, translation initiation region and first 61 codons of the *Lactobacillus amylovorus* α-amylase gene, was inserted between the *Bgl*II-*Sph*I sites of pGTC3. The first 38 amino acids of the primary translation product (preproprotein) comprise the typical tripartite structure of a Sec-dependent signal sequence. The vector constructed in this way was called pGAM3 (Figure 2B-b).

iv) To obtain a P_{amy}-directed expression system with its regulatable expression characteristics, but without secretion directing signals, the *Pst*I-*Sph*I sequence (236 bp) was replaced by a synthetic oligonucleotide with the same restriction enzyme termini. This oligonucleotide contained the entire sequence from *Pst*I to the translation start codon and the first 7 amino acid codons of the *amy* gene. The shortened sequence comprises the AT-rich 5' region but lacks the signal sequence peptidase site. The vector constructed in this way was called pGAMΔss (Figure 2B-c).

Thirdly, the terminator sequence of the *Lactobacillus casei ldh* gene was inserted in the vectors. Between the *Bam*HI and *Bst*EII sites of the three promoter-terminator (P_{ldh}, P_{ldh}-ss_{nprP}, P_{amy}-Δss and P_{amy}) expression vectors, a synthetic oligonucleotide was inserted encompassing the transcription terminator of the *Lactobacillus casei ldh* gene flanked by two *Not*I sites and followed by a nucleotide sequence containing *Bam*HI and *Nco*I sites. Insertion of this oligonucleotide destroyed the original *Bam*HI site but left the *Bst*EII site intact. In the three vectors obtained in this way the promoter sequences are separated from the terminator of the *cbh* gene by an additional terminator. ORF-encoding DNA fragments, to be expressed in *Lactobacillus* but not to be expressed in *Escherichia coli*, can be cloned in the region between the two terminator sequences, preventing transcription (Figure 3).

Fourth, the *E. coli uidA* gene coding for β-glucuronidase was isolated

from pNOM2 [Roberts et al, Curr. Genet. 15, 177-180 1989] and cloned into a pUC19 derivative. The *uidA* gene was extended with two codons by insertion of a hexanucleotide comprising an *Xho*I site (Figure 4). A *Bst*EII site was introduced 6 nucleotides downstream from the translation stop codon. All changes, which were introduced by PCR, were confirmed by sequence analysis. An 1.8-kb *Nco*I-*Bst*EII fragment from this pUC-*uidA* plasmid with the modified *uidA* gene was cloned into the expression vector which was cut with *Nco*I and *Bst*EII, between the *ldh* and *cbh* terminators (Figure 5-b). The presence of unique *Xho*I and *Bst*EII sites facilitates insertion of additional sequences.

Fifth, the series of vectors was further extended by insertion of a PCR-generated *Xho*I-*Bst*EII fragment of the *prtP* gene from *Lactobacillus casei* ATCC 393 encoding 117 C-terminal amino acids (Figure 5-c). This C-terminal sequence is, in analogy with similar sequences from lactococcal proteases, thought to be responsible for anchorage to the bacterial cell surface. The vectors can be further adapted to accomodate differences in thickness of the peptidoglycan layers of different bacterial species by extension of the cell wall spanning domain. The cell wall spanning peptide can be easily extended by replacement of the present *Xho*I-*Bst*EII fragment by a larger *Xho*I-*Bst*EII fragment encoding additional aminoacid sequences, or insertion into the *Xho*I site of a sequence encoding additional amino acids of the cell wall spanning domain. Alternatively, by using a smaller *Xho*I-*Bst*EII fragment (less amino acids in cell wall spanning sequence), the anchored protein will be closer to or in direct contact with the cell surface.

Sixth, a synthetic oligonucleotide comprising restriction enzyme sites for the following enzymes, *Cla*I, *Bam*HI, *Mun*I, *Kpn*I, *Hin*dIII, *Sma*I, *Bbs*I, *Eco*RI, *Sal*I, *Apa*I, *Nco*I (MCS) was inserted between the *Bam*HI and *Nco*I sites of one of the expression cassetttes. The original *Bam*HI site was destroyed using this procedure, but a new *Bam*HI site was present within the newly inserted synthetic oligonucleotide (Figure 6-b). From these prototype cassettes the MCS sequences can easily be transferred to other cassettes as all cassettes contain unique *Bam*HI and *Nco*I sites. With these cassettes any open reading frame (ORF) can be aligned in-frame with that of the N-terminal sequence of the *amy*, *ldh* or *prtP* genes. ORF's that are cloned in one of the sites of the MCS or between two sites of the MCS, can be easily aligned with the *uidA* gene. Alignment of the reading

frame of the new ORF's to sequences downstream from the MCS can be achieved by adjusting the sequence in the distal part of this MCS. For this purpose the *Bbs*I restriction enzyme recognition sequence was incorporated within the synthetic oligonucleotide (MCS). The enzyme *Bbs*I recognizes the non-palindromic sequence GAAGAC and cleaves the DNA strand 2 nucleotides downstream from this sequence. Cleavage in the anti-parallel strand occurs 6 nucleotides downstream from the recognition sequence (Figure 7).

By cutting the MCS (depicted in Figure 6-b,c,d) in one of the expression or expression-secretion cassettes with *Bbs*I, followed by filling in with dNTP and T4 DNA polymerase and ligation, the MCS was enlarged by 4 nucleotides. The reading frame of the sequences upstream of and including the *Bbs*I site shows a +1 frameshift with respect to the methionine codon within the *Nco*I site (Figure 8-b). The thus obtained frameshifted MCS containing expression cassette can again undergo *Bbs*I cutting, filling in with dNTP and ligation. The resulting sequence shows again a +1 frameshift of the sequences upstream of and including the *Bbs*I site, relative to the *Nco*I site (Figure 8-c). The *Bbs*I recognition sequence is not affected by this procedure. The *Bbs*I recognition sequence in the MCS can also be used to clone *Eco*RI compatible sequences in the MCS. The nucleotides for the MCS directly following the *Bbs*I site were chosen to yield compatible termini after *Bbs*I and *Eco*RI cutting. The latter enzyme will not be used in practice since it is not unique in the expression vectors.

The presence of unique *Xho*I and *Bst*EII sites in the expression cassettes allows variation at will of the length of the anchor sequence, so as to adjust the peptidoglycan spanning region to the thickness of the peptidoglycan layer of the host bacterium.

Seventh, the six expression cassettes, digested with *Eco*RI were ligated to pLPE323 [Posno et al, Appl. Envionm. Microbiol. 57, 1822-1828 1991] which was also cut with *Eco*RI, resulting in the following *Escherichia coli-Lactobacillus* shuttle vectors, pLP401-T, pLP402-T, pLP403-T and pLP501-T, pLP502-T and pLP503-T.

Finally, the nucleotide sequence with the terminator, T_{ldh}, was removed from the pLP400-T and pLP500-T series of plasmids by digestion with *Not*I

and religation. The religated plasmids were used to transform *Lactobacillus*, resulting in plasmids pLP401, pLP402, pLP403, and pLP501, pLP502 and pLP503 (Figure 9a and 9b).

Food-grade vectors

The applicability of the vectors can be further improved by removal of the *Escherichia coli* DNA sequences and replacement of the erythromycin-resistance gene by a food-grade selection marker, resulting in a food-grade vector. These alterations can be achieved by digestion of plasmids of the pLP400 and pLP500 series with *Bgl*II and *Nhe*I and insertion of a DNA fragment comprising the *xylR*, *xylA/B* genes and a replicon for *Lactobacillus* (Figure 10). *Lactobacillus casei* or *Lactobacillus plantarum* transformed with vectors harbouring these *xyl* genes can be selected on plates with xylose as sole energy source.

To demonstrate the usefulness of the vectors for the efficient expression of foreign proteins, first the expression of β-glucuronidase was measured. Beta-glucuronidase activity of transformants expressing the *uiA* gene was determined semi-quantitatively in a plate assay using the chromogenic substrate X-gluc. In such assays colonies of *Lactobacillus* transformants that express *uidA* stain blue, whereas colonies of transformants that do not express the *uidA* gene are white. Beta-glucuronidase activity of transformants expressing the *uidA* gene can be quantitatively and reproducibly assayed enzymatically or immunologically using anti-β-glucuronidase antibodies.

The results, which are detailed below in the Examples, show that

i) expression of *uidA* under the control of the *ldh* promoter is constitutive, whereas expression under the control of the *amy* promoter is induced by cellobiose and repressed by glucose,

ii) efficient expression of β-glucuronidase takes place with vectors harbouring the *ldh* promoter or the *amy* promoter,

iii) β-glucuronidase is secreted when the promoter is followed by the α-amylase or proteinase secretion signal,

iv) β-glucuronidase is secreted and remains surface-bound when the *uidA* gene is fused to the anchor sequence of the *prtP* gene.

It is apparent from this and the following examples that expression of a nucleic acid sequence encoding a protein or polypeptide in a lactic acid

bacterium using an expression vector according to the invention in any of the embodiments disclosed and claimed can be improved in comparison to existing systems. It has now become possible to express such a sequence in an amount exceeding 1% of the total amount of protein, preferably exceeding 2% and more preferably even exceeding 3% in an expression host. In particular this is quite spectacular for heterologous nucleic acid expression in a lactic acid bacterium. Most especially this is relevant for expression in *Lactobacillus* a microorganism renown for difficulties for expression of heterologous sequences. The subject invention now enables production of foodgrade lactic acid bacteria comprising foodgrade heterologous sequences. The possibilities for production of proteins and polypeptides that are foodgrade using vectors and methods according to the invention are numerous.

A method for production comprises use of an expression vector and/or host cell comprising such a vector according to the invention in a manner known per se. This involves transformation of a lactic acid bacterium with an expression vector according to the invention comprising the nucleic acid sequence encoding the product of interest. A number of such nucleic acid sequences have been indicated elsewhere in the description. Following introduction of the expression vector in the lactic acid bacterium, the host cell must be cultured in a medium in a manner to be determined in a manner known per se by a person skilled in the art best suited to the kind of host cell, the promoter and/or nucleic acid sequence to be expressed such that expression occurs. Determination of the optimal conditions lies within reach of a person skilled in the art without requiring inventive step or undue experimentation.

The method for constructing an expression vector according to the invention also falls within the scope of protection. Such a method comprises introduction of a vector according to the invention for maintenance and manipulation in a working organism other than the strain of lactic acid bacterium in which expression is to occur, manipulating the vector in a manner known per se to obtain desired characteristics in said working organism and finally introducing the vector comprising a nucleic acid sequence encoding a protein or polypeptide in the strain of lactic acid bacterium in which expression is to occur, said sequence preferably being heterologous to the strain of lactic acid bacterium in which expression is to occur. In particular the method may be useful when

the working organism is a non lactic acid bacterium. In such a method when the vector comprises the transcription terminator sequence located between the promoter and the first codon of the nucleic acid sequence encoding a protein or a polypeptide the vector can be treated in a manner known per se for excising nucleic acid sequences to excise the terminator sequence in order to allow transcription of the nucleic acid sequence encoding a protein or a polypeptide. This can occur e.g. by treatment of the vector with restriction enzyme for cutting the vector at the sites flanking the terminator sequence. Another possibility is that the vector for maintenance is introduced into the expression host prior to the excision and due to the presence of direct repeats flanking the terminator sequence and treatment rendering homologous recombination possible subsequently the expression can be rendered possible due to excision of the terminator sequence in such a fashion.

A vector according to the invention for maintenance and manipulation in a working organism of a nucleic acid sequence encoding a polypeptide or protein, said vector comprising an expression promoter sequence and a cloning site for said nucleic acid sequence encoding a polypeptide or protein with the cloning site being located such that said nucleic acid sequence when present in the vector is or will be controlled by said promoter sequence and is or will be present in an open reading frame in phase with a translation initiation region, is characterised in that a terminator sequence recognisable as such by the working organism in which the vector is to be maintained prior to introduction into a lactic acid bacterium is located on the vector such that the complete polypeptide or protein encoded by the nucleic acid sequence cannot be expressed. Preferably the working organism is a non lactic acid bacterium. A multiple cloning site will be preferable as cloning site.

In another embodiment a vector according to the invention will comprise the nucleic acid sequence encoding the protein or polypeptide at the appropriate cloning site. In such a vector the terminator sequence is present in one particular embodiment between the promoter sequence and the first codon of the nucleic acid sequence encoding a protein or polypeptide. Specifically in a particular embodiment hereof the terminator sequence is present between the promoter sequence and the translation initiation region preceding the nucleic acid sequence. Preferably any of the preceding embodiments may further comprise the

transcription terminator sequence such that it is excisable from the vector in a manner known per se to a person skilled in the art e.g when the transcription terminator sequence is flanked by restriction enzyme sites uniquely present in the vector at such flanking position or is flanked by direct repeats e.g. excisable by homologous recombination by the host.

All the components described for the expression vector according to the invention may naturally also be comprised in the maintenance and manipulation vector according to the invention. Thus such a vector may further comprise one or more of the components:

i) a replicon for a lactic acid bacterium

ii) a DNA sequence encoding the signal sequence of a secreted protein recognisable by a lactic acid bacterium as such this could e.g. be a signal sequence from a lactic acid bacterium, said signal sequence being located in frame between the initiation codon and the nucleic acid sequence encoding a protein or a polypeptide,

iii) a reporter gene encoding an easily assayable product, such a reporter gene e.g. being β-glucuronidase from *Escherichia coli*,

iv) a transcription terminator sequence recognisable by a lactic acid bacterium, said sequence being located downstream of the nucleic acid sequence encoding a protein or polypeptide,

v) a DNA sequence encoding an anchor sequence from a lactic acid bacterium gene and said DNA sequence being located in frame downstream from the nucleic acid sequence encoding a protein or a polypeptide and

vi) a DNA sequence encoding a product with health stimulating properties.

vii) solely foodgrade components.

Example 1

The expression and expression-secretion vectors pLP401.2.3-T and pLP501.-2.3-T, that were constructed in *Escherichia coli*, were digested with *Not*I and ligated, in order to obtain vectors in which the T_{ldh} sequence was omitted. *Lactobacillus casei* ATCC 393 was transformed with the ligation mixtures using electroporation [Posno et al, Appl. Envionm. Microbiol. 57, 1822-1828 (1991)]. Transformants were selected on selective media. Selective media contained 10 g of proteose pepton, 5 g meat extract, 5 g yeast extract, 1 ml Tween 80, 5 g Na-acetate.$3H_2O$, 0.2 g $MgSO_4.7H_2O$, 50 mg $MnSO_4.4H_2O$), 2 g diammoniumcitrate, 17 g agar per 800 ml of deionized water. After autoclaving (20 min at 120ºC) 200 ml 0.5 M potassium-

phosphate buffer pH 7.1 was added. After cooling down to 60ºC, erythromycin and X-gluc (final concentrations 5 and 40µg/ml, respectively) and an energy source (glucose or cellobiose; 0.5-1.0%) were added. The media were poured into petri dishes (13 cm diameter; 25-30 ml medium), after solidification the agar containing plates were dried for 20 min at 60ºC and used immediately. *Lactobacillus* transformants containing vectors from the pLP400 or pLP500 series were streaked onto the surface of the agar plates. Colonies and their phenotypical characteristics (blue or white) could be seen within 24 hrs of incubation at 37ºC.

When transformants of the pLP400 series were plated on indicator plates with X-gluc containing cellobiose as energy source all colonies were light blue, but were white when glucose was used as energy source. When transformants of the pLP500 series were plated on indicator plates all colonies were dark blue, irrespective of whether cellobiose or glucose was used as energy source.

Lactobacillus transformants containing plasmids pLP401, pLP402, pLP403 and with pLP501, pLP502 and pLP503, showed the expected phenotype, whereas *Lactobacillus* transformants containing the pLP401,2,3-T or pLP501,2,3-T plasmids containing the T_{ldh} sequence displayed the white phenotype on selective medium. Reintroduction of pLP503 and pLP402 isolated from *Lactobacillus casei* transformants into *Escherichia coli* DH5α yielded transformants with blue (30%) or white (70%) phenotype on selective medium (L medium, 0.5% glucose, 1.5% agar, 50 µg/ml ampicillin, 40 µg/ml X-gluc). The white phenotype was correlated with structurally instable plasmids (size smaller than starting material). When *Escherichia coli* transformants with a blue phenotype were suspended in physiological salt solution and streaked onto agar plates with selective medium, most colonies showed a white phenotype, indicating plasmid instability.

Example 2

A DNA fragment encoding the SA epitope (Site A; Wiley et al, Nature 89, 373-378 (1981)) of influenza virus comprising 18 amino acids (VTAACS-HAGKSSFYRNLL) was inserted between the *Bam*HI and *Nco*I sites of plasmids pLP401,2,3-T and plasmids pLP501,2,3-T. No structural instability of these vectors was observed in *Escherichia coli*. After removal of the T_{ldh} terminator sequence by digestion with *Not*I and ligation, *Lactobacillus*

casei ATCC 393 was transformed with the resulting plasmids pLP401-SA, pLP402-SA, pLP403-SA and with pLP501-SA, pLP502-SA and pLP503-SA. Transformants were selected on selective media as described in Example 1. When transformants of the pLP400-SA series were plated on indicator plates with X-gluc containing cellobiose as energy source all colonies were light blue, but were white when glucose was used as energy source. When transformants of the pLP500-SA series were plated on indicator plates all colonies were dark blue, irrespective of whether cellobiose or glucose was used as energy source.

Example 3

A DNA fragment encoding the SA epitope of influenza virus comprising 18 amino acids was inserted between the *Bam*HI and *Nco*I sites of derivatives of plasmids pLP401 and pLP501 that lack the *uidA* gene, resulting in plasmids pLP401-SA-1 and pLP501-SA-1. In plasmids pLP401-SA-1 and pLP501-SA-1 the SA encoding sequence is directly fused, in-frame, to the anchor region encoding sequence. *Lactobacillus casei* ATCC 393 was transformed with plasmids pLP401-SA-1 and pLP501-SA-1. Transformants were selected on selective media as described in Example 1. Western analysis shows that transformed bacteria synthesize a product of the expected size that specifically reacts with antibodies against the SA epitope.

Example 4

Amplification of antigenic determinant encoding the SA epitope of influenza virus. An oligonucleotide duplex with *Bam*HI and *Nco*I compatible termini encoding the SA epitope was synthesized (Figure 11) and inserted between the *Bam*HI and *Nco*I sites of plasmids pLP401 and pLP501, resulting in plasmids pLP401-T-SA-1 and pLP501-T-SA-1. No structural instability of these vectors was observed in *Escherichia coli*. Ligation of the *Bam*HI compatible terminus with the *Bam*HI site of plasmids pLP401-T and pLP501-T destroyed the *Bam*HI recognition sequence. Ligation of the *Nco*I compatible terminus with the *Nco*I site of these plasmids left the *Nco*I site intact. The sequence within the synthetic oligonucleotide duplex between the epitope encoding nucleotides and the *Nco*I recognition sequence contains a *Bbs*I site (see Figure 7) which gives rise to a *Bam*HI compatible terminus after cutting with *Bbs*I (Figure 11). Digestion of pLP401-T-SA-1 and pLP501-T-SA-1 with *Bbs*I and *Nco*I permits insertion of a second SA epitope encoding oligonucleotide (pLP401-T-SA-3 and pLP501-T-SA-3). By repeating *Bbs*I and *Nco*I digestion, followed by insertion of an SA epitope encoding

oligonucleotide, expression-secretion vectors (pLP401-T-SA-3 and pLP501-T-SA-3) with 3 copies of the SA epitope sequence fused, in-frame, with the *uidA* gene sequence were obtained.

The procedure described in the previous paragraph allows to easily construct vectors with multiple copies of the same, or different antigenic determinants, provided that the sequence distal to the antigenic determinant encoding sequence contains *Bbs*I (as described in the preceding paragraph) and *Nco*I recognition sequences. Although the basic expression and expression-secretion vectors contain a PRESS with a *Bbs*I site (Figure 6-b,c,d), this *Bbs*I site is removed when antigenic determinant encoding sequences are inserted between the *Bam*HI and *Nco*I sites.

After removal of the T_{ldh} terminator sequence by digestion with *Not*I and ligation, *Lactobacillus casei* ATCC 393 was transformed with the resulting plasmids pLP401-SA-1,2,3 and pLP501-SA-1,2,3. Transformants were selected on selective media as described in Example 1. When transformants of the pLP400-SA series were plated on indicator plates with X-gluc containing cellobiose as energy source all colonies were light blue, but were white when glucose was used as energy source. When transformants of the pLP500-SA series were plated on indicator plates all colonies were dark blue, irrespective of whether cellobiose or glucose was used as energy source. Western analysis shows that transformed bacteria synthesize a product of the expected size that specifically reacts with antibodies against the SA epitope and antibodies against β-glucuronidase.

Example 5

A DNA fragment encoding the Hacket epitope of influenza virus comprising 10 amino acids (FERFEIFPKE; Hacket et al, J. Immunol. 135, 1391-1395 (1985)) was inserted between the *Bam*HI and *Nco*I sites of plasmids pLP401-T, pLP402-T and pLP503-T, resulting in plasmids pLP401-T-HA, pLP402-T-HA and pLP503-T-HA. No structural instability of these vectors was observed in *Escherichia coli*. After removal of the T_{ldh} terminator sequence by digestion with *Not*I and ligation, *Lactobacillus casei* ATCC 393 was transformed with the resulting plasmids pLP401-HA, pLP402-HA and pLP503-HA and *Lactobacillus plantarum* strains ATCC 14917, ATCC 8014, NCIB 8826, 80, NCDO 1193 with pLP402-HA and pLP503-HA. Transformants were selected on selective media as described in Example 1. When transformants of the pLP400-HA series were plated on indicator plates with X-gluc containing

cellobiose as energy source all colonies were light blue, but were white when glucose was used as energy source. When transformants of pLP503-HA were plated on indicator plates all colonies were dark blue, irrespective of whether cellobiose or glucose was used as energy source. Western analysis shows that transformed bacteria synthesize a product of the expected size that specifically reacts with antibodies against the HA epitope and antibodies against β-glucuronidase.

The amount of fusion protein synthesized by pLP503-HA was quantitatively determined by measuring the β-glucuronidase activity and by quantitative ElISA. The enzymatic assay and the ELISA determination indicate that the fusion protein amounts to 1-3% of total cellular protein.

Example 6

A DNA fragment encoding pig rotavirus VP8 polypeptide comprising the first 248 amino acids of the VP4 protein (this part corresponds with the region of the VP4 protein that arises after trypsic cleavage [Lopez et al, 65, 3738-3745 (1991)]) was inserted between the *Bam*HI and *Nco*I sites of plasmids pLP401-T, pLP402-T, pLP501-T and pLP502-T, resulting in plasmids pLP401-T-R08, pLP402-T-R08, pLP501-T-R08 and pLP502-T-R08. No structural instability of these vectors was observed in *Escherichia coli*. After removal of the T_{ldh} terminator sequence by digestion with *Not*I and ligation, *Lactobacillus casei* ATCC 393 was transformed with plasmids pLP401-R08, pLP402-R08, pLP501-R08, and pLP502-R08. Transformants were selected on selective media as described in Example 1. When transformants of the pLP400-R08 series are plated on indicator plates with X-gluc containing cellobiose as energy source all colonies are light blue, but are white when glucose is used as energy source. When transformants of the pLP500-R08 series are plated on indicator plates all colonies are dark blue, irrespective of whether cellobiose or glucose is used as energy source.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1

Schematic representation of the construction route of some embodiments of the expression vectors.

a. Structure of plasmid pGTC3. T_{cbh}, transcription terminator of *Lactobacillus plantarum cbh* gene.

b. Nucleotide sequence comprising PRESS, poly restriction enzyme site sequence, and part of the sequence containing the *cbh* terminator.

Figure 2: Schematic representation of the construction route of the expression vectors.

Figure 3: Sequence of the synthetic oligonucleotide duplex encompassing the transcription terminator of the L-*ldh* gene of *Lactobacillus casei*. Inverted repeats capable of forming a stem loop structure are marked by horizontal arrows. The overhanging terminal sequences on the left and right side of the sequences are compatible to the *Bam*HI and *Bst*EII sites, respectively, of the promoter-terminator cassette vectors.

Figure 4

Nucleotide sequence of the 3'end of the *uidA* gene before (a) and after (b) mutagenesis by PCR. An *Xho*I site was introduced immediately upstream of the translation stop codon to allow in-frame fusion of amino acid encoding sequences of the *uidA* gene with an ORF to be fused with the 3' end of this *uidA* gene.

Figure 5

a. Physical map of promoter-double-terminator cassette vector pPTT3. P stands for P_{ldh}, P_{ldh}-ss_{prtP}, P_{amy}-Δss, or P_{amy}, T_{ldh} and T_{cbh}: transcription terminators of the *Lactobacillus casei ldh* and *Lactobacillus plantarum cbh* gene, respectively.

b. Physical map of promoter-terminator cassette in which an adjusted *Escherichia coli uidA* gene was cloned.

c. As b. The reading frame of the *uidA* gene is fused to that of the 3' end of the *prtP* gene coding for the anchor sequence.

Figure 6

Nucleotide sequences of the region between the *ldh* gene transcription terminator sequence and the *uidA* gene of the expression cassette vectors.

a. Nucleotide sequence between the *Bam*HI and *Nco*I sites in pGTT3; this sequence is part of the synthetic oligonucleotide sequence encompassing the *Lactobacillus casei ldh* transcription terminator signal.

b. Nucleotide sequence of MCS inserted between *Bam*HI and *Nco*I sites of pPTUT. The original *Bam*HI was destroyed using this strategy but a new *Bam*HI site is present in the MCS. Starting from this MCS two new,

slightly modified MCS were obtained in which the *Bam*HI site and following sites are frame-shifted.

c. A vector containing the MCS as depicted in b) was digested with *Cla*I and the resulting DNA termini were filled in with dNTP's using T4 DNA polymerase. After ligation the MCS contained a *Bam*HI site shifted two nucleotides with respect to *G*.

d. Nucleotide sequence of a MCS obtained by replacing the *Cla*I-*Bam*HI sequence in b) by a *Cla*I-*Bsp*EI-*Bam*HI sequence. The *Bam*HI site is shifted four nucleotides with respect to *G*.

DNA fragments can be cloned in MCS depicted in b, c or d depending on the reading frame needed. To ensure alignment of the cloned ORF with the *uidA* gene, restriction enzyme sites just upstream of the *Nco*I site (containing translation startcodon ATG of the *uidA* gene) can be modified with other restriction enzymes (see elsewhere in the description).

Figure 7

Cleavage pattern of *Bbs*I. The bold hexanucleotide duplex represents the *Bbs*I recognition sequence, the arrows show the cleavage sites. N and N' stand for any nucleotide and its complementary counterpart.

Figure 8

a) Nucleotide sequence of MCS sequences as depicted in Figure 6-b,c,d.

b) Nucleotide sequence of MCS after filling up the *Bbs*I restriction enzyme generated termini of MCS depicted in a) with dNTP's using T4 polymerase and ligation.

c) Nucleotide sequence of MCS after filling up the *Bbs*I restriction enzyme generated termini of MCS depicted in b) with dNTP's using T4 polymerase and ligation. The line between G-G corresponds with the nucleotide sequences of Figure 6-b,c,d.

Figure 9

a. Schematic representation of expression vector.

b. Schematic structure of cassettes present in plasmids of the pLP400 and pLP500 series. P_{ldh} and P_{amy} represent the promoters of the *ldh* gene of *Lactobacillus casei* and *amy* promoter of the α-amylase gene of *Lactobacillus amylovorus*, respectively. *uidA*, β-glucuronidase gene of *Escherichia coli*. T_{ldh}, transcription terminator of the *ldh* gene of *Lactobacillus casei*. T_{cbh}, transcription terminator of the *cbh* gene of *Lactobacillus*

plantarum. Anchor *prtP*, sequence from *Lactobacillus casei* encoding the membrane anchor sequence and peptidoglycan spanning region. The versatility of the vectors was further improved by introduction, between the *Bam*HI and *Nco*I sites, of a multicloning region (MCS) comprising the following restriction enzyme sites, *Cla*I, *Bam*HI, *Mun*I, *Kpn*I, *Hin*dIII, *Sma*I, *Bbs*I, *Eco*RI, *Sal*I, *Apa*I, *Nco*I.

In the promoter-terminator (P_{ldh}, P_{ldh}-ss_{prtP}, P_{amy}-Δss and P_{amy}) expression vectors, the *Bam*HI site, and all other sites of the MCS, are present in three reading frames, allowing in-frame translation of any DNA sequence fused to the N-terminal sequence of LDH, α-amylase or PrtP. The presence of a unique cloning sites in the MCS permits easy alignment, if necessary, of the reading frames of the antigenic determinant encoding sequence and of *uidA*.

Figure 10
Schematic representation of the structure of part of the xylose cluster of genes of *Lactobacillus pentosus*. Replacement of the erythromycin-resistance marker by these genes allows a transformed *Lactobacillus* strain to grow on xylose as sole energy source.

Figure 11
Nucleotide sequence of the distal part of the MCS containing the *Nco*I and *Bbs*I recognition sites. The recognition sequences for *Bam*HI, *Nco*I and *Bbs*I are indicated with bold letters.

Figure 12
The nucleotide sequence of S layer protein in particular the promoter region of *L. acidophilus slpA*.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Nederlandse Organisatie....
(B) STREET: Schoemakerstraat 97
(C) CITY: Delft
(D) STATE: Zuidholland.
(E) COUNTRY: The Netherlands
(F) POSTAL CODE (ZIP): NL-2628 VK

(ii) TITLE OF INVENTION: A novel method for the construction of vectors for lactic acid bacteria like lactobacillus.....

(iii) NUMBER OF SEQUENCES: 16

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 77 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: fragment with cbh terminator of L plantarum

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAATTCAGAT CTACTAGTCT CGAGGCATGC GGATCCATTC AAGGTCACCC CTAGAGTCAA 60

CTAAAAGCCA CTACTGT 77

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 211 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: fragment with l-ldh terminator of L casei

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GATCTAGCGG CCGCCTAGAA CAGCCGGGCA ACTGCCCGGT TGTTCTTTTT TTAGCGGCCG 60

CTAAGGATCC AGGTGAGCTC GGGGTTGCCA TGGCTACTGT CAACGATCGC CGGCGGATCT 120

TGTCGGCCCG TTGACGGGCC AACAAGAAAA AAATCGCCGG CGATTCCTAG GTCCACTCGA 180

GCCCCAACGG TACCGATGAC AGCTGCCAGT G 211

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: fragment with 3'end of uidA gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGCAAATGAA TCACA 15

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: fragment with 3'end of uidA gene fusion to ORF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GGCAAACTCG AGTAATGTAA GGTCACCTCG ACAAGCTT 38

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: BamhI-NcoI fragment of pGTT3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGATCCAGGT GAGCTCGGGG TTGCCATGG 29

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 98 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: BamhI-NcoI fragment of pPTUT with mcs

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GGATCTTATC GATTTACGGA TCCAACAATT GCAAGGTACC GATCAAGCTT ACCCGGGAAG 60

AAGACAGAAT TCACAAGTCG ACCGGGCCCT TGCCATGG 98

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: BamhI-NcoI fragment of pPTUT with mcs
frameshifted 2 nts

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGATCTTATC GCGATTTACG GATCCAACAA TTG 33

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: BamhI-NcoI fragment of pPTUT with mcs frameshifted 4 nts

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGATCTTATC GATGTCCGGA TGGATCCAAC AATTG 35

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: mcs nucleotide sequence of seq 6,7 and 8

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGAAGAAGAC AGAATTCACA AGTCGACCGG GCCCTTGCCA TGG 43

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: mcs amino sequence of seq 9

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Arg Arg Gln Asn Ser Gln Val Asp Arg Ala Leu Ala Met
1 5 10

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: mcs nucleotide sequence of seq 9 with filled BbsI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGAAGAAGAC AGAATTAATT CACAAGTCGA CCGGGCCCTT GCCATGG 47

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: mcs amino sequence of seq 11

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Glu Glu Asp Arg Ile Asn Ser Gln Val Asp Arg Ala Leu Ala Met
1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: mcs nucleotide sequence of seq 6,7 and 8 between first two G's

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGAAGAAGAC AGAATTAATT AATTCACAAG TCGACCGGGC CCTTGCCATG G 51

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: mcs amino sequence of seq 13

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Lys Lys Thr Glu Leu Ile Asn Ser Gln Val Asp Arg Ala Leu Ala Met
1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 382 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: L acidophilus promoter region of SlpA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TGCTTGTGGG GTAAGCGGTA GGTGAAATAT TACAAATAGT ATTTTTCGGT CATTTTAACT 60

TGCTATTTCT TGAAGAGGTT AGTACAATAT GAATCGTGGT AAGTAATAGG ACGTGCTTCA 120

GGCGTGTTGC CTGTACGCAT GCTGATTCTT CAGCAAGACT ACTACCTCAT GAGAGTTATA 180

GACTCATGGA TCTTGCTTTG AAGGGTTTTG TACATTATAG GCTCCTATCA CATGCTGAAC 240

CTATGGCCTA TTACATTTTT TTATATTTCA AGGAGGAAAA GACCACATGA AGAAAAATTT 300

AAGAATCGTT AGCGCTGCTG CTGCTGCTTT ACTTGCTGTT GCTCCAGTTG CTGCTTCTGC 360

TGTATCTACT GTTAGCGCTG CT 382

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:
(A) ORGANISM: L acidophilus N terminal region of SlpA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Met Lys Lys Asn Leu Arg Ile Val Ser Ala Ala Ala Ala Ala Leu Leu
1 5 10 15

Ala Val Ala Pro Val Ala Ala Ser Ala Val Ser Thr Val Ser Ala Ala
20 25 30

Claims

1. Expression vector for expression and optionally for excretion and optionally for surface display of a protein or polypeptide in a lactic acid bacterium host cell, said vector comprising an expression promoter sequence and a nucleic acid sequence encoding said protein or polypeptide with expression of said nucleic acid sequence being controlled by said promoter sequence, characterised in that said encoding nucleic acid sequence is preceded by a 5' non-translated nucleic acid sequence comprising at least the minimal sequence required for ribosome recognition and RNA stabilisation, followed by a translation initiation codon, said translation initiation codon being immediately followed by a fragment of at least 5 codons of the 5' terminal part of the translated nucleic acid sequence of a gene of a lactic acid bacterium or a structural or functional equivalent of said fragment, said fragment also being controlled by said promoter, said fragment not normally being part of the gene comprising the nucleic acid sequence encoding said protein or polypeptide, with the proviso that the fragment does not encode the secretion signals of Lactobacillus Surface Protein, Lactococcus Major Secretion Product and with the proviso that the promoter is not a T7 polymerase promoter, a Lactobacillus Surface Protein promoter or a Lactococcus Major Secretion Product promoter and with the proviso that the vector is not plasmid pKTH1797, pKTH1798, pKTH1799, pKTH1801, pKTH1805, pKTH1806, pKTH1807, pKTH1809 or pKTH1889.

2. Expression vector according to claim 1, wherein the lactic acid bacterium is a *Lactobacillus* or Bifidobacterium.

3. Expression vector according to claim 1 or 2, wherein the lactic acid bacterium is a *Lactobacillus*.

4. Expression vector for expression and optionally for excretion and optionally for surface display of a protein or polypeptide in a lactic acid bacterium host cell, said vector comprising an expression promoter sequence and a nucleic acid sequence encoding said protein or polypeptide with expression of said nucleic acid sequence being controlled by said promoter sequence, characterised in that said encoding nucleic acid sequence is preceded by a 5' non-translated nucleic acid sequence comprising at least the minimal sequence required for ribosome recognition and RNA stabilisation, followed by a translation initiation

codon, said translation initiation codon being immediately followed by a fragment of at least 5 codons of the 5' terminal part of the translated nucleic acid sequence of a gene of a Lactobacillus or a structural or functional equivalent of said fragment, said fragment also being controlled by said promoter, said fragment not normally being part of the gene comprising the nucleic acid sequence encoding said protein or polypeptide with the proviso that the fragment does not encode the secretion signal of Lactobacillus Surface Protein and with the proviso that the promoter is not a Lactobacillus Surface Protein promoter.

5. Expression vector according to any of the preceding claims, wherein the 5' terminal part of the translated nucleic acid sequence is richer in A-T than the rest of said gene such that a person skilled in the art will consider it an A-T rich sequence.

6. Expression vector according to any of the preceding claims wherein said 5' terminal part of the translated nucleic acid sequence has an A-T content of 60% or more, preferably 62% or more, preferably 65% or more, suitably 70% or more.

7. Expression vector according to claim 3 wherein said 5' terminal part of the translated nucleic acid sequence has an A-T content of 100% or less, preferably 95% or less, preferably 90% or less, suitably 85% or less.

8. Expression vector according to any of the preceding claims, wherein the 5' non translated nucleic acid sequence is homologous to a lactic acid bacterium, preferably homologous to the lactic acid bacterium host cell.

9. Expression vector according to any of the preceding claims, wherein the 5' non translated nucleic acid sequence is the complete 5' non translated region including the Shine Dalgarno sequence.

10. Expression vector according to any of the preceding claims, wherein the 5' terminal part of the translated nucleic acid sequence comprises a sequence smaller than the complete encoding sequence, preferably smaller than the 5' half of the encoding sequence, more preferably less than 35 codons of the 5' terminal part, most preferably

less than 16 codons of the 5' terminal part.

11. Expression vector according to any of the preceding claims, wherein the 5' non-translated nucleic acid sequence and the 5' terminal part of the translated nucleic acid sequence are derivable from the same gene.

12. Expression vector according to any of the preceding claims, wherein the 5' non-translated nucleic acid sequence and/or the 5' terminal part of the translated nucleic acid sequence and the promoter are derivable from the same gene.

13. Expression vector according to any of the preceding claims, wherein the promoter is homologous to a lactic acid bacterium, preferably homologous to the lactic acid bacterium host cell.

14. Expression vector according to any of the preceding claims, wherein the nucleic acid sequence encoding the protein or polypeptide encodes an antigenic determinant.

15. Expression vector according to any of the preceding claims, wherein the protein or polypeptide is expressed to a degree at least equivalent to that of 1%, preferably more than 2% with more preference for more than 3% of total protein of the lactic acid bacterium host cell.

16. Expression vector according to any of the preceding claims, wherein the protein or polypeptide to be expressed is encoded by the following *Lactobacillus* genes: lactate dehydrogenase encoding gene, the proteinase encoding gene, the amylase encoding gene, the S layer protein encoding gene, the ribosomal RNA encoding gene, the glyceraldehyde 3 phosphate dehydrogenase encoding gene.

17. Expression vector according to any of the preceding claims, wherein the nucleic acid sequence encoding the protein or polypeptide to be expressed in a lactic acid bacterium is heterologous or homologous to a lactic acid bacterium host cell.

18. Expression vector according to any of the preceding claims, wherein the nucleic acid sequence encoding the protein or polypeptide to

be expressed in a lactic acid bacterium is heterologous to a lactic acid bacterium.

19. Expression vector according to any of the preceding claims, wherein the nucleic acid sequence encoding the protein or polypeptide to be expressed in a lactic acid bacterium is heterologous to the lactic acid bacterium host cell.

20. Host cell comprising an expression vector according to any of the preceding claims.

21 Lactic acid bacterium host cell comprising an expression vector according to any of claims 1-19.

22. *Lactobacillus* bacterium or a *Bifidobacterium* host cell comprising an expression vector according to any of claims 1-19.

23. *Lactobacillus* bacterium host cell comprising an expression vector according to any of claims 1-19.

24. Cell according to any of claims 20-23, wherein host cell and expression vector are food grade.

25. Vector for maintenance and manipulation of a nucleic acid sequence encoding a polypeptide or protein in a working organism, said vector comprising an expression promoter sequence and a cloning site for said nucleic acid sequence encoding a polypeptide or protein, said cloning site being located such that said nucleic acid sequence when present in the vector is or will be controlled by said promoter sequence and is or will be present in an open reading frame in phase with a translation initiation region, said vector being characterised in that a terminator sequence recognisable as such by the working organism in which the vector is to be maintained prior to introduction into a lactic acid bacterium is located on the vector such that the complete polypeptide or protein encoded by the nucleic acid sequence cannot be expressed.

26. A vector according to claim 25, wherein said cloning site is preferably a multiple cloning site.

27. A vector according to claim 25 or 26, wherein said working organism is preferably a non lactic acid bacterium.

28. Vector according to any of claims 25-27, wherein the nucleic acid sequence encoding the protein or polypeptide is present in the vector at the appropriate cloning site.

29. Vector according to any of claims 25-28, wherein the terminator sequence is present between the promoter sequence and the first codon of the nucleic acid sequence encoding a protein or polypeptide.

30. Vector according to any of claims 25-29, wherein the terminator sequence is present between the promoter sequence and the translation initiation region preceding the nucleic acid sequence.

31. Vector according to any of claims 25-30, wherein the transcription terminator sequence is present such that it is excisable from the vector in a manner known per se e.g. when the transcription terminator sequence is flanked by restriction enzyme sites uniquely present in the vector at such flanking position or is flanked by direct repeats e.g. excisable by homologous recombination by the host.

32. Vector according to any of claims 25-31, wherein furthermore the components of an expression vector according to any of claims 1-19 are present.

33. Vector according to any of claims 25-31, wherein furthermore the components of an expression vector for expression and optionally for excretion and optionally for surface display of a protein or polypeptide in a lactic acid bacterium host cell are present, said expression vector comprising an expression promoter sequence and a nucleic acid sequence encoding said protein or polypeptide with expression of said nucleic acid sequence being controlled by said promoter sequence, characterised in that said encoding nucleic acid sequence is preceded by a 5' non-translated nucleic acid sequence comprising at least the minimal sequence required for ribosome recognition and RNA stabilisation, followed by a translation initiation codon, said translation initiation codon being immediately followed by a fragment of at least 5 codons of the 5' terminal part of the translated nucleic acid sequence of a gene of a

lactic acid bacterium or a structural or functional equivalent of said fragment, said fragment also being controlled by said promoter, said fragment not normally being part of the gene comprising the nucleic acid sequence encoding said protein or polypeptide.

34. Vector according to any of the claims 1-19 and 25-33 further comprising one or more of the components:
i) a replicon for a lactic acid bacterium
ii) a DNA sequence encoding the signal sequence of a secreted protein recognisable by a lactic acid bacterium as such, said signal sequence being located in frame between the initiation codon and the nucleic acid sequence encoding a protein or a polypeptide,
iii) a reporter gene encoding an easily assayable product, such a reporter gene e.g. being β-glucuronidase from *Escherichia coli*,
iv) a transcription terminator sequence recognisable by a lactic acid bacterium, said sequence being located downstream of the nucleic acid sequence encoding a protein or polypeptide,
v) a DNA sequence encoding an anchor sequence from a lactic acid bacterium gene and said DNA sequence being located in frame downstream from the nucleic acid sequence encoding a protein or a polypeptide and
vi) a DNA sequence encoding a product with health stimulating properties.

35. Vector according to any of the claims 1-19 and 25-34, wherein the vector is a foodgrade vector.

36. Method for obtaining improved stability and degree of expression of a nucleic acid sequence heterologous to a lactic acid bacterium, said heterologous sequence being present on a vector according to any of claims 1-19 as such or in combination with any of claims 32-35, said vector being introduced into a lactic acid bacterium in a manner known per se, said expression of heterologous nucleic acid sequence occurring in the lactic acid bacterium.

37. Method for constructing a vector according to any of claims 1-19 as such or in combination with any of claims 32-35, comprising introducing a vector according to any of claims 25-35 into a working organism other than the strain of lactic acid bacterium in which expression is to occur, manipulating the vector in a manner known per se to obtain desired characteristics in said working organism and finally

introducing the vector comprising nucleic acid sequence encoding a protein or polypeptide in the strain of lactic acid bacterium in which expression is to occur, said sequence preferably being heterologous to the strain of lactic acid bacterium in which expression is to occur.

38. Method according to claim 37, wherein the working organism is a non lactic acid bacterium.

39. Method according to claim 37 or 38 wherein the transcription terminator sequence located between the promoter and the first codon of the nucleic acid sequence encoding a protein or a polypeptide is treated in a manner known per se for excising nucleic acid sequences to excise the terminator sequence in order to allow transcription of the nucleic acid sequence encoding a protein or a polypeptide e.g. by treatment of the vector with restriction enzyme for cutting the vector at the sites flanking the terminator sequence.

40. A method for producing a protein or polypeptide heterologous to a lactic acid bacterium in a lactic acid bacterium, preferably a *Lactobacillus* or *Bifidobacterium* in an amount exceeding 1%, preferably exceeding 2% of the total amount of protein, said method comprising expression of nucleic acid encoding such a protein or polypeptide from an expression vector according to any of claims 1-19 as such or in combination with 32-35 or a host cell according to any of claims 20-24 in a manner known per se.

41. Protein or polypeptide obtainable by the method of claim 40.

42. Protein or polypeptide according to claim 41, wherein the protein or polypeptide is foodgrade.

Fig. 1a

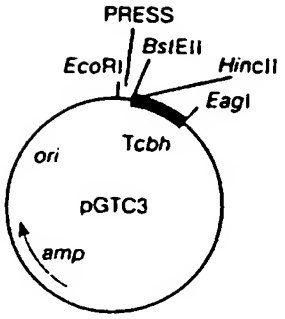


Fig. 1b.

EcoRI BglII SpeI XhoI SphI BamHI
GAATTC AGATCT ACTAGT CTCGAG GCATGC GGATCC ATTCAA

BstEII HincII
GGTCACC CCTAGA GTCAAC TAAAAGCCACTACTGT
N stop ---->
T_{cbh} sequences

Fig. 2A

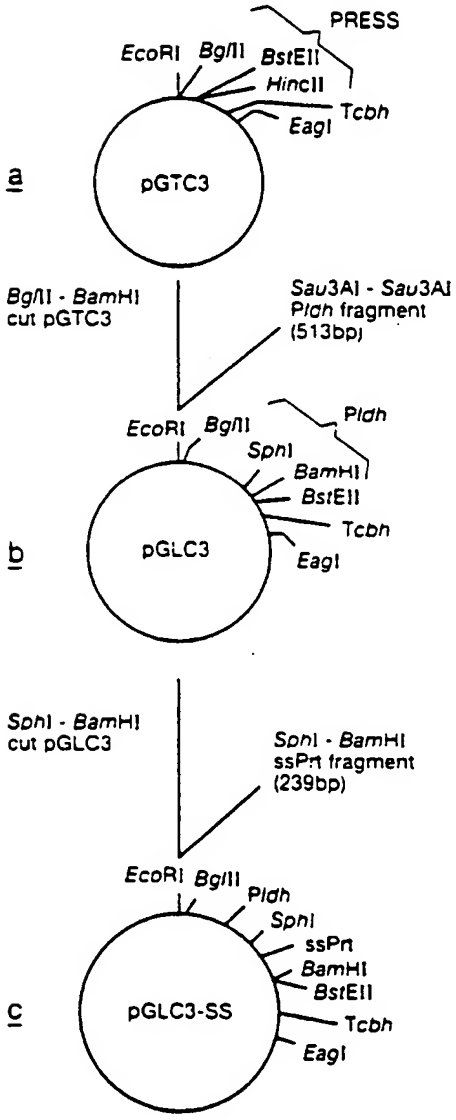

PRESS
EcoRI
BglII
BstEII
HincII
Tcbh
EagI
a
pGTC3
BglII - BamHI
cut pGTC3
Sau3AI - Sau3AI
Pldh fragment
(513bp)
Pldh
EcoRI
BglII
SphI
BamHI
BstEII
Tcbh
b
pGLC3
EagI
SphI - BamHI
cut pGLC3
SphI - BamHI
ssPrt fragment
(239bp)
EcoRI
BglII
Pldh
SphI
ssPrt
BamHI
BstEII
Tcbh
c
pGLC3-SS
EagI

Fig. 2B

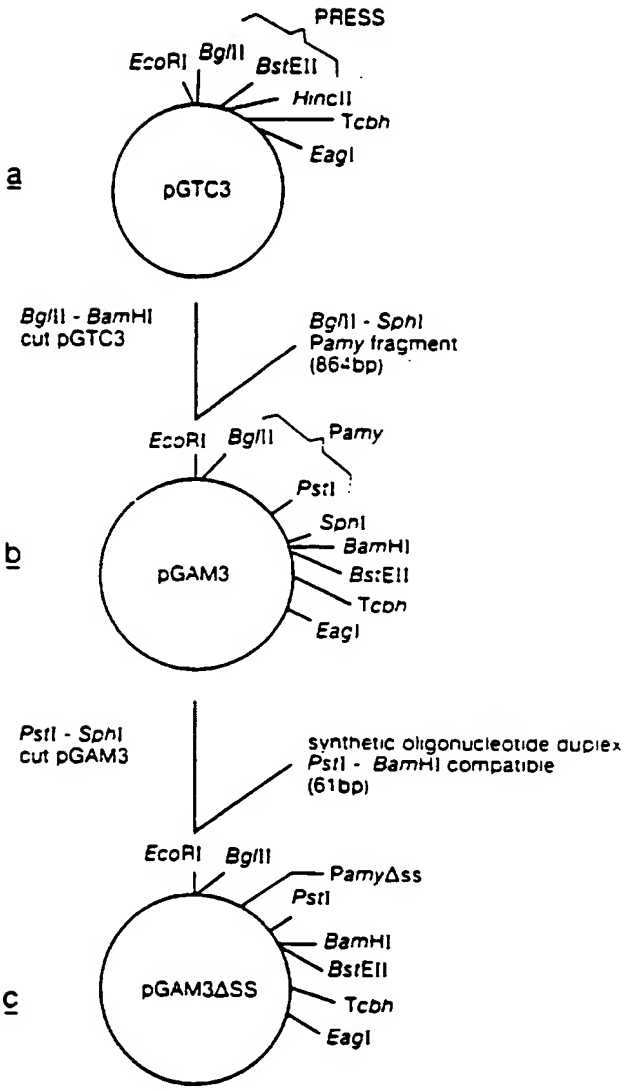

a
PRESS
EcoRI
BglII
BstEII
HincII
TcbH
EagI
pGTC3
BglII - BamHI
cut pGTC3
BglII - SphI
Pamy fragment
(864bp)
b
EcoRI
BglII
Pamy
PstI
SphI
BamHI
BstEII
TcbH
EagI
pGAM3
PstI - SphI
cut pGAM3
synthetic oligonucleotide duplex
PstI - BamHI compatible
(61bp)
c
EcoRI
BglII
PamyΔss
PstI
BamHI
BstEII
TcbH
EagI
pGAM3ΔSS

Fig. 3

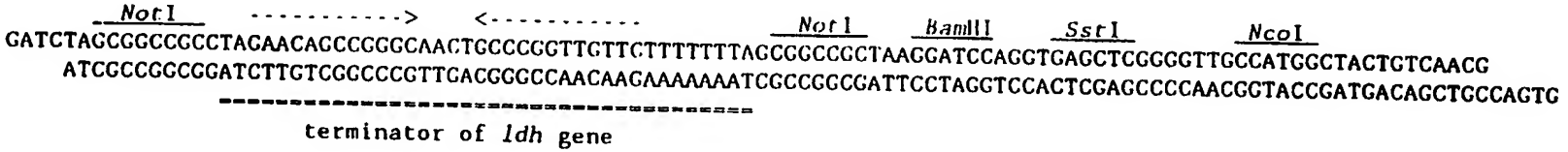


Fig. 4

Nucleotide sequence of the 3'end of the *uidA* gene

a. GGC.AAA.TGAATCACA
G K stop

b. GGC.AAA.CTC.GAG.TAATGTAACGTCACCTCGACAAGCTT
XhoI BstEII HindIII
G K L E stop pUC cloning site

Fig. 5

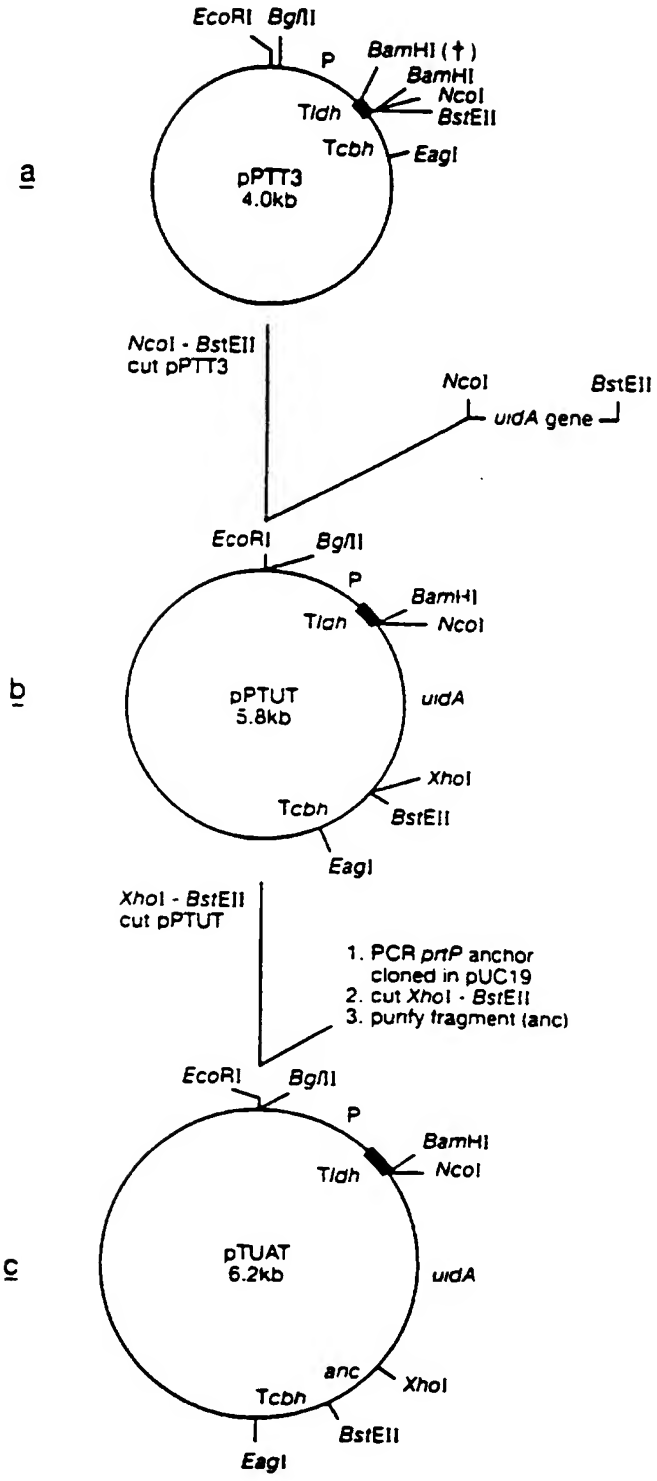

a
EcoRI
BglII
P
BamHI (†)
BamHI
NcoI
BstEII
Tldh
Tcbh
EagI
pPTT3
4.0kb
NcoI - BstEII
cut pPTT3
NcoI
BstEII
uidA gene
b
EcoRI
BglII
P
BamHI
NcoI
Tldh
pPTUT
5.8kb
uidA
XhoI
Tcbh
BstEII
EagI
XhoI - BstEII
cut pPTUT
1. PCR prtP anchor
cloned in pUC19
2. cut XhoI - BstEII
3. purify fragment (anc)
c
EcoRI
BglII
P
BamHI
NcoI
Tldh
pTUAT
6.2kb
uidA
anc
XhoI
Tcbh
BstEII
EagI

Fig. 6

a. BamHI SstI NcoI
G GAT CCA GGT GAG CTC GGG GTT GCC ATG G
D P G E L G V A M

b. ClaI BamHI MunI KpnI HindIII
G GAT CTT ATC GAT TTA CGG ATC CAA CAA TTG CAA GGT ACC GAT CAA GCT T
D L I D L R I Q Q

SmaI BbsI EcoRI SalI ApaI NcoI
ACC CGG GAA GAA GAC AGA ATT CAC AAG TCG ACC GGG CCC TTG CCA TGG

c. NruI BamHI MunI
G GAT CTT ATC GCG ATT TAC GGA TCC AAC AAT TG ---> sequence as in B
D L I A I Y G S N

d. ClaI BspEI BamHI MunI
G GAT CTT ATC GAT GTC CGG ATG GAT CCA ACA ATT G ---> sequence as in B
D L I D V R M D P

Fig. 7

↓
-----GAAGAC N N N N N N N N N-----
-----CTTCTG N'N' N'N'N'N' N'N'N'-----
↑

Fig. 8

a.

```
                BbsI  EcoRI             SalI    ApaI        NcoI
G-------GA AGA AGA CAG AAT TCA CAA GTC GAC CGC GCC CTT GCC ATG G
            R   R   Q   N   S   Q   V   D   R   A   L   A   M
```

b.

```
           BbsI
G-------GAA GAA GAC AGA ATT AAT TCA CAA GTC GAC CGG GCC CTT GCC ATG G
         E   E   D   R   I   N   S   Q   V   D   R   A   L   A   M
```

c.

```
            BbsI
G-------G AAG AAG ACA GAA TTA ATT AAT TCA CAA GTC GAC CGG GCC CTT GCC ATG G
           K   K   T   E   L   I   N   S   Q   V   D   R   A   L   A   M
```

Fig. 9a

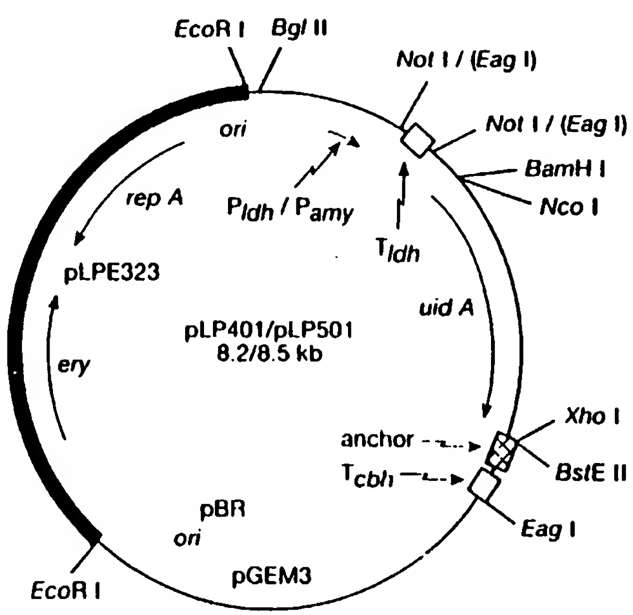


Fig.9b

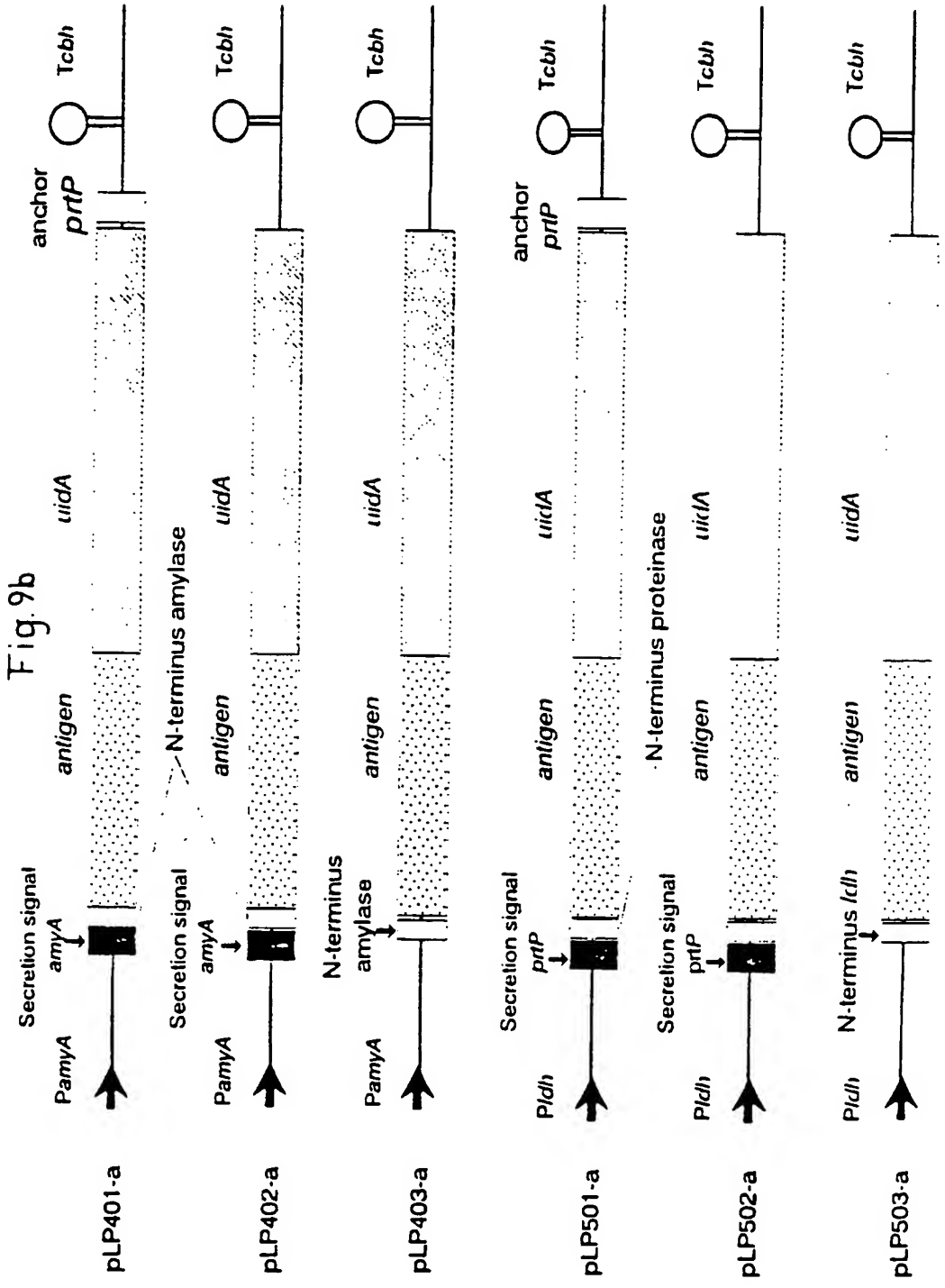

pLP401-a
PamyA
Secretion signal
amyA
antigen
uidA
anchor
prtP
Tcbh
pLP402-a
N-terminus amylase
pLP403-a
N-terminus
amylase
pLP501-a
PldhL
prtP
pLP502-a
N-terminus proteinase
pLP503-a
N-terminus ldh

Fig.10

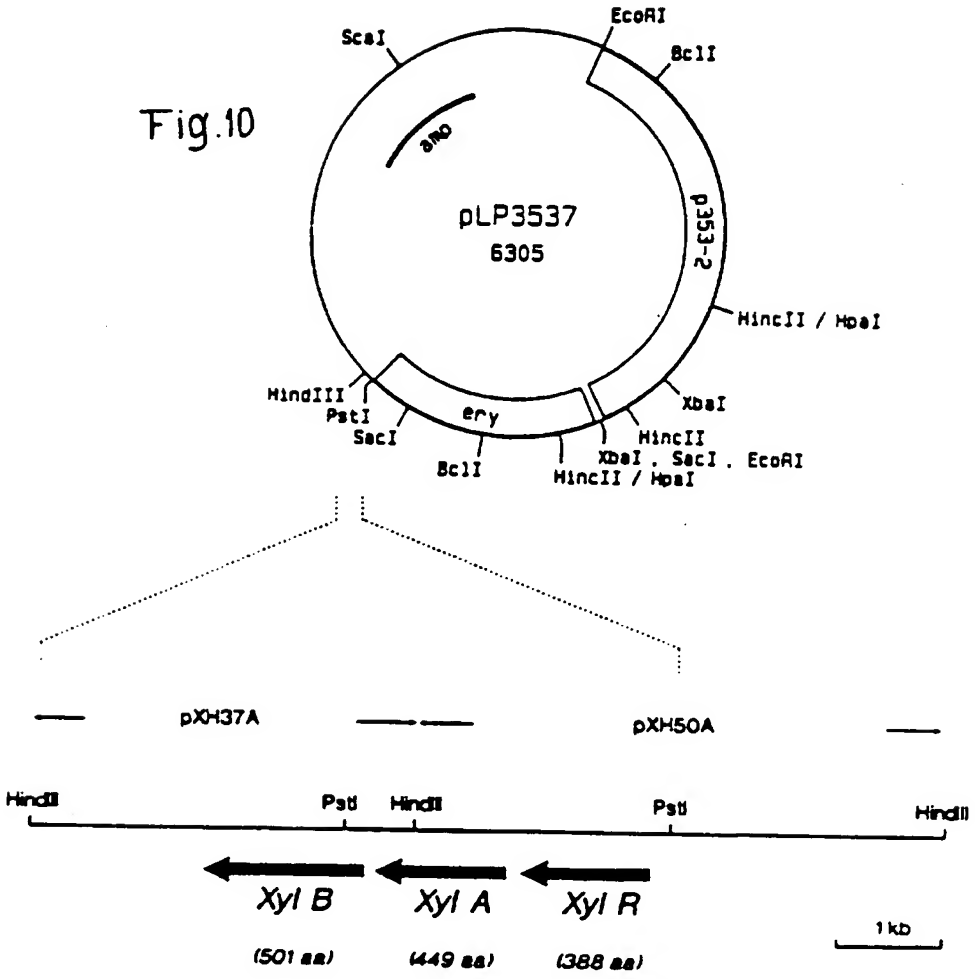

ScaI
EcoRI
BclI
amp
pLP3537
6305
p353-2
HincII / HpaI
HindIII
PstI
SacI
ery
BclI
HincII / HpaI
XbaI
HincII
XbaI . SacI . EcoRI
pXH37A
pXH50A
HindII
PstI
HindII
PstI
HindII
Xyl B
Xyl A
Xyl R
(501 aa)
(449 aa)
(388 aa)
1 kb

Fig.11

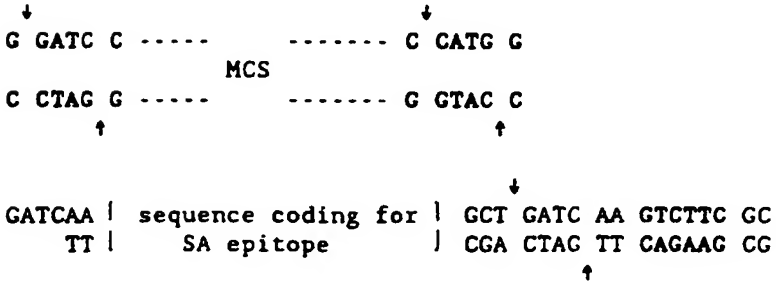
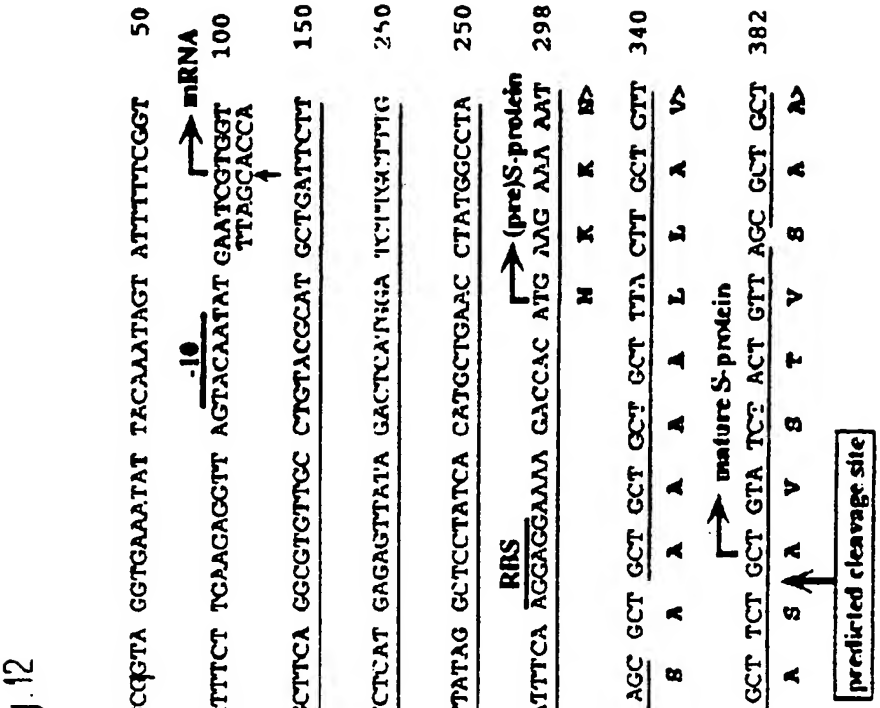
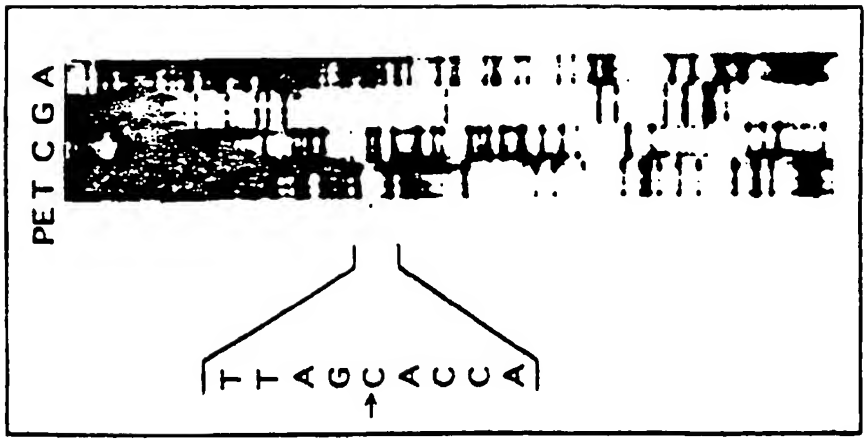

G GATC C -----
C CTAG G -----
MCS
C GATC C -----
G CTAG G -----
GATCAA
TT
sequence coding for
SA epitope
GCT GATC AA GTCTTC GC
CGA CTAG TT CAGAAG CG

Fig.12


TGCTTGTGGG GTAAGCCGTA GGTGAAATAT TACAAATAGT ATTTTTCGGT 50
-35
CATTTAACT TGCTATTTCT TGAAGAGGTT AGTACAATAT -10 GAATCGTGGT 100
mRNA
TTAGCACCA
AAGTAATAGG ACGTGCTTCA GGCGTGTTGC CTGTACGCAT GCTGATTCTT 150
CAGCAAGACT ACTACCTCAT GAGAGTTATA GACTCATTGA TCTTTATTTG 200
AAGGGTTTTG TACATTATAG GCTCCTATCA CATGCTGAAC CTATGGCCTA 250
RBS
TTACATTTTT TTATATTTCA AGGAGGAAAA GACCAC ATG (pre)S-protein AAG AAA AAT 298
M K K N
TTA AGA ATC GTT AGC GCT GCT GCT GCT TTA CTT GCT GTT 340
L R I V S A A A A L L A V
mature S-protein
GCT CCA GTT GCT GCT TCT GCT GTA TCT ACT GTT AGC GCT GCT 382
A P V A A S A V S T V S A A
predicted cleavage site


PETCGA
TTAGCACCA

INTERNATIONAL SEARCH REPORT

Interr al Application No
PCT/NL 96/00160

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/74 C12N15/62 C12N15/46 C12N15/44 C12N15/70
C12N15/31 C07K14/14 C07K14/11 //C12N1/21,C12R1:225

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FEMS MICROBIOL. REVIEWS, vol. 88, 1992, ELSEVIER, AMSTERDAM, NL, pages 73-92, XP000573739 M. VAN DE GUCHTE ET AL.: "Gene expression in Lactococcus lactis" see the whole document --- -/--	1-4,8,9, 11-14, 16-24, 41,42

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
27 June 1996	18.07.96
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Authorized officer Hornig, H

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KEYSTONE SYMPOSIUM ON MUCOSAL IMMUNITY: NEW STRATEGIES FOR PROTECTION AGAINST VIRAL AND BACTERIAL PATHOGENS, KEYSTONE, COLORADO, USA, JANUARY 16-23, 1995. JOURNAL OF CELLULAR BIOCHEMISTRY SUPPLEMENT 0 (19A). 1995. 255. ISSN: 0733-1959, ABSTRACT NO. J1-203, XP002006853 BOERSMA W J A ET AL: "Lactobacillus as vectors with intrinsic adjuvanticity for safe live mucosal vaccines." see abstract	1-4,8,9, 11-14, 16-24, 34,41,42
X	EP,A,0 564 965 (NESTLE SA) 13 October 1993 see the whole document	1-4,8,9, 11-14, 16-24, 41,42
X	WO,A,93 17117 (LYNXVALE LTD) 2 September 1993 cited in the application	4,8,9, 11-14, 17-24, 41,42
Y	see the whole document	1-3
X	WO,A,92 04451 (GENESIT OY ;VALIO FINNISH CO OPERATIVE DAI (FI)) 19 March 1992 cited in the application	4,8,9, 11-14, 17-24, 41,42
Y	see the whole document	1-3
X	EP,A,0 449 770 (CIBA GEIGY AG) 2 October 1991 cited in the application	4,8,9, 11-14, 17-24, 41,42
Y	see the whole document	1-3
X	US,A,5 242 821 (PALVA ILKKA ET AL) 7 September 1993 cited in the application	4,8,9, 11-14, 17-24, 41,42
Y	see the whole document	1-3
X	WO,A,94 00581 (VIAGEN OY) 6 January 1994 cited in the application	41,42
Y	see the whole document	1-3

-/--

INTERNATIONAL SEARCH REPORT

Interr al Application No
PCT/NL 96/00160

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	J. BIOTECHNOL. (1996), 44(1-3), 183-92 CODEN: JBITD4;ISSN: 0168-1656, January 1996, XP000572655 POUWELS, PETER H. ET AL: "The potential of Lactobacillus as a carrier for oral immunization: Development and preliminary characterization of vector systems for targeted delivery of antigens" see the whole document -----	1-42

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Interr ıal Application No
PCT/NL 96/00160

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0564965	13-10-93	EP-A-	0569604	18-11-93
		AU-B-	665626	11-01-96
		AU-B-	3670993	14-10-93
		BR-A-	9301468	13-10-93
		CA-A-	2093152	08-10-93
		JP-A-	6054692	01-03-94
		NZ-A-	247338	28-08-95
		US-A-	5491079	13-02-96
WO-A-9317117	02-09-93	EP-A-	0628083	14-12-94
		GB-A,B	2278358	30-11-94
		JP-T-	7504815	01-06-95
WO-A-9204451	19-03-92	AU-B-	652920	15-09-94
		EP-A-	0550428	14-07-93
		JP-T-	6503945	12-05-94
EP-A-0449770	02-10-91	AU-B-	643511	18-11-93
		AU-B-	7355891	03-10-91
		CA-A-	2038706	23-09-91
		JP-A-	4211384	03-08-92
US-A-5242821	07-09-93	CA-A-	1338137	12-03-96
WO-A-9400581	06-01-94	AU-B-	4329793	24-01-94